Vol. XXXVI Parts III and IV December 1949

BIOMETRIKA

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

MANAGING EDITOR

E. S. PEARSON

ASSOCIATE EDITORS

M. G. KENDALL JOHN WISHART

in consultation with

HARALD CRAMÉR R. C. GEARY

J. B. S. HALDANE

Reprinted by offset-litho, 1960

ISSUED BY

THE BIOMETRIKA OFFICE, UNIVERSITY COLLEGE, LONDON

PRINTED AT THE UNIVERSITY PRESS, CAMBRIDGE

[*Issued* 23 *January* 1950]

VOLUME XXXVI, PARTS III AND IV DECEMBER 1949

THE ESTIMATION OF THE PARAMETERS OF TOLERANCE DISTRIBUTIONS

BY D. J. FINNEY

Lecturer in the Design and Analysis of Scientific Experiment, University of Oxford

1. INTRODUCTION

In many types of investigation concerning the distribution of tolerances in a population, direct measurement of tolerance for each individual in a sample is impossible; the only practicable procedure may be to subject individuals to specified 'doses' of the stimulus under test and to record whether or not they show a particular response. For example, in the evaluation of the toxic properties of insecticides, a different dose of an insecticide may be given to each of several batches of insects and the numbers of deaths and survivals recorded. Insects killed by a certain dose are known to have a tolerance lower than that dose, insects surviving are known to have a higher tolerance, but only the one test can be made on each insect and the parameters of the distribution of tolerance can therefore be estimated only by comparison of the proportions of responses recorded for different doses. Other examples of data of this kind have been given elsewhere (Finney, 1947*a*).

A method of analysis based upon the probit transformation is now commonly used to assist the interpretation of quantal response data of the kind just described. The writer has published (1947*a*) a detailed account of the standard probit technique for estimating the parameters of a normal distribution of tolerance. The purpose of the present paper is to derive and illustrate a general method, applicable to tolerance distributions other than the normal, which includes various modifications in the specification of the problem such as those already discussed by Finney (1944, 1947*b*) and Wadley (1949).

2. THE EQUIVALENT DEVIATE TRANSFORMATION

Suppose that each member of a population has a tolerance, u, in respect of some stimulus, and that the distribution of u depends only upon a parameter of location, μ, and a parameter of scale, σ. The distribution may then be written

$$\frac{1}{\sigma} f\left(\frac{u-\mu}{\sigma}\right) du, \tag{1}$$

where $f(v)$ is a function containing no unknown parameters and

$$\int_{-\infty}^{\infty} f(v)\, dv = 1. \tag{2}$$

Either or both of the tails of the distribution may have finite limits, but for convenience these cases may be regarded as contained within the infinite limits without affecting the general theory; the limits of the distribution, of course, are assumed to be independent of the parameters. The probability that a stimulus whose measure on the same scale as u is x will produce the characteristic response in an individual chosen at random is

$$P = \int_{-\infty}^{x} \frac{1}{\sigma} f\left(\frac{u-\mu}{\sigma}\right) du. \tag{3}$$

For example, in a population of insects whose tolerance of a certain insecticide is distributed normally, the probability that a dose x would kill a particular insect is

$$P = \int_{-\infty}^{x} \frac{1}{\sqrt{(2\pi)\sigma}} \exp\left[-\frac{(u-\mu)^2}{2\sigma^2}\right] du.$$

The units for u and x need not be those in which the experimenter habitually measures, and often a dose metameter, such as the logarithm of the concentration of insecticide, is required in order that the distribution shall take the right form.

Write now

$$v = \frac{u-\mu}{\sigma} = \alpha + \beta u \text{ say,} \tag{4}$$

where

$$\alpha = -\mu/\sigma, \quad \beta = 1/\sigma. \tag{5}$$

Then the general expression for P may be written

$$P = \int_{-\infty}^{\alpha+\beta x} f(v)\, dv. \tag{6}$$

For any specified function $f(v)$, a quantity Y, the *equivalent deviate* of P, may be defined as a monotonic increasing function of P by the equation

$$P = \int_{-\infty}^{Y} f(v)\, dv. \tag{7}$$

In particular, for a normal distribution of tolerances

$$f(v) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}v^2},$$

and Y is the *normal equivalent deviate* of P, differing from the probit of P only by omission of the conventional addition of 5. Equations (6) and (7) then show that the relationship between the probability of response and the measure of the stimulus, which is equivalent to a statement of the tolerance distribution, may be expressed by

$$Y = \alpha + \beta x. \tag{8}$$

Consequently, the procedure for estimating the parameters μ and σ may be regarded as finding a linear relationship between x and the Y-transform of the probability of response; this probability is itself estimated from experiments by p, the proportion of individuals responding to a particular dose, and equations (5) enable estimates of μ and σ to be obtained from α and β. When the proportion responding to a particular dose can be directly observed, in an experiment in which a known number of individuals are given the dose and those responding to it are counted, the sequence of calculations for solving the maximum likelihood estimation equations is the analogue of the standard probit calculations corresponding to the form of tolerance distribution adopted. The theory of this, with examples of its consequences for several tolerance distributions, has been given elsewhere (Finney, 1947*a*, *b*).

3. The estimation of the parameters

Experimentation leading to direct observation of values of p for given values of x is not always possible. For example, the population may contain an unknown proportion of individuals who will respond whatever dose is given, or the total number of individuals receiving a particular dose may have to be estimated from a parallel sample instead of being counted. Methods of analysis for these two problems have been described (Finney, 1944; Wadley,

1949), and their similarity suggests that they are particular instances of some more general theorem. Such a theorem will now be derived.

Suppose that the tolerance distribution for individuals is that given above as (1), so that the probability of response at dose x *due to the action of that dose* is P ($= 1-Q$, say). Suppose further that an experiment is performed in which r, the number of individuals responding, can be observed for a series of values of the dose, x, and that, under the conditions of experiment, the probability of exactly r responses at a particular x is

$$P(r) = F(J+KQ); \tag{9}$$

here J and K are additional parameters which in general will also require estimation, and $F(\ )$ is a known function. For example, if a proportion C_1 of the population will respond even to zero dose and a proportion C_2 is unable to respond however large the dose, then the probability that an individual selected at random will respond when it receives a dose x is

$$P^* = (1-C_2)-Q(1-C_1-C_2). \tag{10}$$

If n individuals are subjected to a dose x, the number of responses will have a binomial distribution based upon this probability:

$$P(r) = \binom{n}{r} P^{*r} Q^{*n-r}, \tag{11}$$

which is of the form of equation (9). Unless C_1 and C_2 are known, direct estimation of P from tests at a particular x is impossible, and the whole data must be used in order to estimate also the additional parameters. If C_2 is known to be zero, for a normal distribution of tolerances this reduces to the problem of taking account of 'natural mortality' in probit analysis (Finney, 1944, 1947*a*, Chapter 6).

The logarithm of the likelihood of the results obtained for a series of doses may be written

$$L = S \log P(r) = S \log F, \tag{12}$$

where the summation denoted by S may include doses for which $x \to -\infty$, for which Q is known to be unity, or $x \to +\infty$, for which Q is known to be zero. If both J and K are unknown, the maximum likelihood estimates of α, β, J, K will be obtained by equating to zero the partial differential coefficients of L with respect to each of these; if, as sometimes happens, the conditions of the problem specify J or K exactly, the corresponding equation will be omitted.

Now
$$\frac{\partial Q}{\partial \alpha} = -f(v) = -Z, \text{ say, and } \frac{\partial Q}{\partial \beta} = -xZ, \tag{13}$$

where Z is the ordinate of the standardized tolerance distribution, $f(v)\,dv$. Consequently the maximum likelihood equations may be written

$$\left.\begin{aligned} 0 = \frac{\partial L}{\partial \alpha} &= -KS\frac{ZF'}{F}, \quad 0 = \frac{\partial L}{\partial J} = S\frac{F'}{F}, \\ 0 = \frac{\partial L}{\partial \beta} &= -KS\frac{xZF'}{F}, \quad 0 = \frac{\partial L}{\partial K} = S\frac{QF'}{F}, \end{aligned}\right\} \tag{14}$$

where
$$F'(\lambda) = \frac{\partial}{\partial \lambda}\{F(\lambda)\}. \tag{15}$$

Equations (14) can be solved most easily by an iterative process which calculates adjustments to provisional values of α, β, J, K. The adjustments may be calculated from Taylor-Maclaurin expansions to the first order, in the usual manner (cf. Finney, 1947*a*, Appendix II,

1947*b*). If J, K are approximations to the estimates of these parameters obtained at the end of one cycle of calculations, an empirical rate of response to one of the doses tested may be defined as $p\ (=1-q)$, the maximum likelihood estimate of P from the data for that dose alone when J and K have these values. Thus q is the solution of

$$F'(J+Kq)=0. \tag{16}$$

Hence, to the first order in $(Q-q)$,

$$F'(J+KQ)=K(Q-q)\,F''(J+KQ). \tag{17}$$

The second differential coefficients of L, after substitution of the expected values of all parameters, are

$$\left.\begin{aligned}
&\frac{\partial^2 L}{\partial\alpha^2}=K^2S\frac{Z^2F''}{F}, && \frac{\partial^2 L}{\partial J^2}=S\frac{F''}{F},\\
&\frac{\partial^2 L}{\partial\alpha\partial\beta}=K^2S\frac{xZ^2F''}{F}, && \frac{\partial^2 L}{\partial\alpha\partial K}=-KS\frac{QZF''}{F},\\
&\frac{\partial^2 L}{\partial\beta^2}=K^2S\frac{x^2Z^2F''}{F}, && \frac{\partial^2 L}{\partial\beta\partial K}=-KS\frac{xQZF''}{F},\\
&\frac{\partial^2 L}{\partial\alpha\partial J}=-KS\frac{ZF''}{F}, && \frac{\partial^2 L}{\partial J\partial K}=S\frac{QF''}{F},\\
&\frac{\partial^2 L}{\partial\beta\partial J}=-KS\frac{xZF''}{F}, && \frac{\partial^2 L}{\partial K^2}=S\frac{Q^2F''}{F}.
\end{aligned}\right\} \tag{18}$$

Introduce now a *weight*, W, defined by

$$W=-\frac{K^2Z^2F''}{F}, \tag{19}$$

and two auxiliary variates, t_1, t_2, defined by

$$t_1=-\frac{1}{KZ},\quad t_2=-\frac{Q}{KZ}. \tag{20}$$

If α, β, J, K are approximations to the maximum likelihood estimates, obtained by a graphical method, by rough calculation, or from a previous cycle of the calculations about to be described, the quantities defined by equations (19) and (20) may be formed for each dose, with the aid of equations (6), (9) and (13). Adjustments to the estimates, $\delta\alpha$, $\delta\beta$, δJ, δK are then the solution of a set of linear equations whose coefficients are given by equations (18); these equations may be written:†

$$\left.\begin{aligned}
\delta\alpha SW+\delta\beta SWx+\delta J SWt_1+\delta K SWt_2&=SW\left(\frac{Q-q}{Z}\right),\\
\delta\alpha SWx+\delta\beta SWx^2+\delta J SWxt_1+\delta K SWxt_2&=SWx\left(\frac{Q-q}{Z}\right),\\
\delta\alpha SWt_1+\delta\beta SWxt_1+\delta J SWt_1^2+\delta K SWt_1t_2&=SWt_1\left(\frac{Q-q}{Z}\right),\\
\delta\alpha SWt_2+\delta\beta SWxt_2+\delta J SWt_1t_2+\delta K SWt_2^2&=SWt_2\left(\frac{Q-q}{Z}\right).
\end{aligned}\right\} \tag{21}$$

† If the approximations to J and K differ much from the maximum likelihood estimates, some values of p may be negative or greater than unity. Since p, q are introduced only as aids to the solution of the maximum likelihood equations, these values are nevertheless to be used in all that follows.

The *working equivalent deviate* (analogous to the working probit), y, must now be introduced. It is defined as

$$y = Y + \frac{Q-q}{Z}, \tag{22}$$

where Y, Q, Z, q are determined from the approximations to the estimates of the parameters by means of equations (8), (7), (13) and (16). The introduction of the symbol $\overline{\theta}$ for the weighted mean of a variate θ,

$$\overline{\theta} = SW\theta/SW, \tag{23}$$

and $S_{\theta\phi}$ for the weighted sum of products of deviations of any two variates θ, ϕ,

$$\begin{aligned} S_{\theta\phi} &= SW(\theta-\overline{\theta})(\phi-\overline{\phi}) \\ &= SW\theta\phi - (SW\theta)(SW\phi)/SW, \end{aligned} \tag{24}$$

enables the equations, after some rearrangement, to be reduced to

$$\alpha_1 = \overline{y} - \beta_1\overline{x} - \delta J\overline{t}_1 - \delta K\overline{t}_2, \tag{25}$$

and

$$\left.\begin{aligned} \beta_1 S_{xx} + \delta J S_{xt_1} + \delta K S_{xt_2} &= S_{xy}, \\ \beta_1 S_{xt_1} + \delta J S_{t_1t_1} + \delta K S_{t_1t_2} &= S_{t_1y}, \\ \beta_1 S_{xt_2} + \delta J S_{t_1t_2} + \delta K S_{t_2t_2} &= S_{t_2y}. \end{aligned}\right\} \tag{26}$$

Here α_1, β_1 have been written for the adjusted values $(\alpha+\delta\alpha), (\beta+\delta\beta)$, but in general this is unnecessary as there is little danger of confusion. Equations (26) show that β_1, δJ, δK are the weighted partial regression coefficients of y on x, t_1, t_2 respectively.

Providing that no doses are included for which P is known absolutely, instead of merely estimated, equations (26) are complete as shown. Often, however, 'control' batches of subjects are tested in the absence of the stimulus. If the dose scale is a simple measure of weight or concentration, this will correspond to $x = 0$ in a tolerance distribution whose tail does not extend below $x = 0$; if a logarithmic dose metameter is being used, the control batches have $x \to -\infty$. For these controls, $Q = 1$ is known, and therefore they make no contribution to information on α and β; it does not follow that q, derived from equation (16), is unity, and the controls do give information on J and K. Examination of the method of derivation of equations (26) shows that $(-F''/F)$, evaluated for $Q = 1$, must then be added to $S_{t_1t_1}$, $S_{t_1t_2}$ and $S_{t_2t_2}$, and that $K(1-q)\,F''/F$ must be added to S_{t_1y} and S_{t_2y}. The other coefficients of the equations remain unchanged. Less commonly, a maximal dose $(x \to +\infty)$ is given, for which the conditions of the problem require $Q = 0$ exactly. From equation (9), it is clear that the subjects so tested can give information only on J, which is therefore in some way a measure of that part of the population of subjects immune to the stimulus. Only two coefficients of equations (26) are affected: $(-F''/F)$ must be added to $S_{t_1t_1}$ and $K(1-q)\,F''/F$ must be added to S_{t_1y}.

After the solution of the equations, a second cycle of calculation may be performed, using $\alpha_1, \beta_1, J_1\,(= J+\delta J), K_1\,(= K+\delta K)$ in place of α, β, J, K, and iteration may continue as long as seems desirable. Of course, in the limit δJ, δK become zero and β becomes simply the weighted regression coefficient of y on x, but for practical purposes the iteration need seldom be carried to the stage at which the adjustments to J and K are negligible.

If k doses in all have been tested (including any for which $x \to -\infty$ or $x \to +\infty$), a χ^2 with $(k-4)$ degrees of freedom will test whether the observations agree satisfactorily with expectations based upon the parameters. This χ^2 can be calculated by comparison of observed

and expected frequencies. A simpler calculation, correct when the maximum likelihood estimates are used throughout, is

$$\chi^2_{[k-4]} = S_{yy} - \beta S_{xy}, \tag{27}$$

and even when the iteration is not carried to the limit a satisfactory approximation will usually be given by

$$\chi^2_{[k-4]} = S_{yy} - \beta S_{xy} - \delta J S_{t_1 y} - \delta K S_{t_2 y}. \tag{28}$$

The χ^2 calculated by either of these formulae corresponds to the usual expression

$$\chi^2 = S \frac{(\text{observed-expected})^2}{\text{expected}}$$

taken over all classes, and consequently does not allow for the undue influence of very small expected frequencies in some classes; this point has been discussed elsewhere (Finney, 1947*a*), and the same method of dealing with it (calculation of each expectation and grouping of classes) is required here. Equation (28), however, is often good enough and saves the more laborious calculations previously used (Finney, 1944; Wadley, 1949), except when borderline values indicate the need for more detailed examination.

For data showing no heterogeneity (i.e. a non-significant χ^2), the variances of the estimates of the parameters may be found in the usual manner. In fact

$$V(\bar{y}) = 1/SW \tag{29}$$

and the variances and covariances of β, J, K are the elements of the inverse matrix

$$V = \begin{pmatrix} S_{xx} & S_{xt_1} & S_{xt_2} \\ S_{xt_1} & S_{t_1 t_1} & S_{t_1 t_2} \\ S_{xt_2} & S_{t_1 t_2} & S_{t_2 t_2} \end{pmatrix}^{-1}; \tag{30}$$

V will generally be calculated as a stage in the solution of equations (26). The mean $\bar{y}$ is uncorrelated with β, J, K, and the variances of α and of other combinations of the parameters can readily be derived. If there is evidence of heterogeneity, but not of a kind that makes the form of the tolerance distribution or some other characteristic of the analysis inappropriate, all variances must be multiplied by the heterogeneity factor; no new discussion of this point is needed (Finney, 1947*a*, §§ 18, 19).

If the conditions of the problem specify J or K, δJ or δK will be everywhere zero, the second or third equation of (26) will be omitted, and χ^2 will have $(k-3)$ degrees of freedom. If both J and K are specified, only the first of equations (26) remains, χ^2 has $(k-2)$ degrees of freedom, and the calculations are those for the standard probit method, or for the corresponding method based on some non-normal tolerance distribution, just as described in a previous paper (Finney, 1947*b*).

4. Applications and illustrations

(i) *Adjustment for natural mortality*

In the testing of insecticides, 'control' or untreated batches of insects frequently show some deaths during the time that elapses between treatment and classification of the treated batches. It may then be assumed that some of the deaths amongst treated insects would have occurred even if the insects have been untreated; they are due to natural causes and not to the insecticide. If an insect has a probability C of natural death, independently of the insecticide,

a dose of insecticide, sufficient to give a probability of death P amongst insects not dying from natural causes, will have associated with it a total death-rate obtained as the particular case of equation (10) with $C_2 = 0$:

$$P^* = 1 - Q(1-C). \tag{31}$$

This assumes that the two causes of death operate independently of one another.†

If n insects all receive the same dose, and react to it independently of one another, the probability that r of them die is

$$P(r) = \binom{n}{r} P^{*r}(1-P^*)^{n-r} = F\{1-Q(1-C)\}, \tag{32}$$

where, in the notation of § 3,

$$J = 1, \quad K = -(1-C). \tag{33}$$

Since J is specified by the conditions of the problem, the second equation of (26) is not required. Now

$$\log F = \text{const.} + r \log P^* + (n-r)\log(1-P^*).$$

Differentiation twice with respect to the argument of $F(\ )$, followed by the replacement of r, $(n-r)$ by their expectations nP^*, $n(1-P^*)$ gives the expected value

$$\frac{F''}{F} = -\frac{n}{P^*(1-P^*)},$$

whence, using equations (19) and (31),

$$W = \frac{nZ^2(1-C)}{Q\{1-Q(1-C)\}}. \tag{34}$$

This may be regarded as equivalent to a *weighting coefficient*, or weight per individual, w, where

$$w = \frac{Z^2}{Q\left\{\frac{1}{1-C} - Q\right\}}, \tag{35}$$

to be multiplied by n in order to give the total weight for the batch. When $C = 0$, w reduces to the familiar formula for probit analysis; the present result, however, applies to tolerance distributions other than the normal providing that the appropriate Q and Z are used.

Direct estimation of C rather than of $(1-C)$ is convenient, and for that purpose the auxiliary variate t_2 of equation (20) is modified to

$$t = Q/Z; \tag{36}$$

t_1, of course, is not required since J is known. The factor $1/K$ is included in the auxiliary variate for theoretical work, but in computation it is most easily left as a multiplier of δJ, δK, of which account may be taken later. Equations (26) now become

$$\beta S_{xx} + \frac{\delta C}{1-C} S_{xt} = S_{xy}, \quad \beta S_{xt} + \frac{\delta C}{1-C} S_{tt} = S_{ty}, \tag{37}$$

and from equation (25),

$$\alpha = \bar{y} - \beta\bar{x} - \frac{\delta C}{1-C}\bar{t}. \tag{38}$$

† Horsfall (1945) has strongly criticized this assumption and any attempts to base methods of statistical analysis upon it. Extensive experimental studies will be needed before discussion of this matter is taken further; in the absence of these, equation (31) is a reasonable mathematical model to try, and experience has proved it to be adequate for the explanation of many sets of data.

Thus β and $\delta C/(1-C)$ are the partial regression coefficients of y on x and t, and when they have been found the revised estimate of the remaining parameter, α, is given by equation (38). The S_{xx}, etc., are weighted sums of squares and products of deviations, the weight for any dose being nw. If some insects, say n_c in all, have been included in the experiment but kept as untreated controls in order to provide a direct estimate of C, contributions from these will affect S_{tt} and S_{ty}. Writing r_c as the number of responses amongst these controls and

$$c = r_c/n_c \tag{39}$$

as the estimate of C based on the controls alone, the comments below equations (26) show that S_{tt} must be increased by

$$\frac{n_c(1-C)}{C}. \tag{40}$$

Since the value of q for the controls is obtained from equation (31) by substitution of $P^* = c$ to give

$$q = \frac{1-c}{1-C},$$

S_{ty} must be increased by

$$\frac{n_c(c-C)}{C}. \tag{41}$$

The calculations based on equations (37)–(41) are equivalent to those previously given (Finney, 1944, 1947*a*); the latter, however, were developed only for a normal distribution of tolerances, but the present theory shows the same form of analysis to be applicable to any distribution providing that the appropriate relationship between Y and P (equation (7)) is employed. The method of calculation now recommended is simpler than the earlier version, especially as most of it follows the standard pattern for multiple regression.

The provisional C used in any cycle of calculations may by chance be greater than some of the p^* for low doses; for each of these, equation (16) will then give q greater than unity, p negative. The corresponding P, being an expected value, is of course positive, and application of the formula for the working equivalent deviate (equation (22)) will enable the results for these doses to exert their proper influence in the calculations even though diagrammatic representation in the usual probit or other regression diagram is impossible.

For a normal distribution of tolerances, μ and σ are respectively the mean and standard deviation. Values of the auxiliary variate, t, defined by equation (36), and of the weighting coefficient, w, from equation (35), have been tabulated by Finney (1947*a*, Table II). As an illustration of the calculations, the example given in Chapter 6 of that book may be analysed again; this is of interest as showing how the natural mortality adjustment can be employed in data for a biological assay, in which two insecticides or other preparations are tested simultaneously.

Preparations of two derris roots, *W*.213 and *W*.214 were tested on the grain beetle, *Oryzaephilus surinamensis*. The log concentrations, in mg. dry root per litre, are shown as x in the first column of Table 1. The columns, n, r, p^* show numbers of insects tested, numbers affected (dead + moribund + slightly affected), and percentages affected respectively. In a control batch of 129 insects, 21 were affected, so that

$$c = 21/129 = 16{\cdot}3\,\%.$$

Inspection of all the data suggested $C = 17{\cdot}0\,\%$

as a first approximation to the required estimate; values of p were then calculated from equation (31) which may be rewritten

$$p = \frac{p^* - 0{\cdot}170}{1 - 0{\cdot}170}.$$

In accordance with usual practice, these have been tabulated as percentages, instead of proportions, though in all formulae p is used as a proportion.

The empirical probits of p (Finney, 1947*a*, Table I) were next tabulated, and, either by inspection of Table 1 or by plotting these probits against x, a column of expected probits, Y, was obtained from two parallel regression lines. The weighting coefficients, w, and values of t for each Y were read from tables (Finney, 1947*a*, Table II), in order to give the next two columns in Table 1; for example, for $Y = 7{\cdot}6$ and $C = 17\,\%$, $w = 0{\cdot}03298$ and multiplication by 142 gives 4·7 as the total weight for the first batch of beetles. The working probits, y, were also read from tables (Finney, 1947*a*, Table IV).

Calculation and summation of columns for nwx, nwt, nwy gave totals from which $\bar{x}$, $\bar{t}$, $\bar{y}$ were formed. Further multiplications by x, t, y led to sums of squares and products of deviations, found according to equation (24). All these were formed separately for the two roots. In order to form single estimates of β and C, corresponding sums of squares and products were added, together with the special contributions to S_{tt} and S_{ty} from the controls obtained from formulae (39)–(41); details are shown at the bottom of Table 1. Equations (37) then take the form

$$28{\cdot}8295\beta - 76{\cdot}5485\frac{\delta C}{1-C} = 80{\cdot}709, \quad -76{\cdot}5485\beta + 897{\cdot}1128\frac{\delta C}{1-C} = -214{\cdot}659.$$

The inverse matrix of coefficients (equation (30)) is easily found as

$$V = \begin{pmatrix} 0{\cdot}0448475 & 0{\cdot}0038267 \\ 0{\cdot}0038267 & 0{\cdot}0014412 \end{pmatrix}. \tag{42}$$

Therefore
$$\beta = 80{\cdot}709 \times 0{\cdot}0448475 - 214{\cdot}659 \times 0{\cdot}0038267$$
$$= 2{\cdot}7982,$$

and
$$\frac{\delta C}{1-C} = 80{\cdot}709 \times 0{\cdot}0038267 - 214{\cdot}659 \times 0{\cdot}0014412$$
$$= -0{\cdot}00052.$$

Hence, by substitution of the provisional value of C,

$$\delta C = -0{\cdot}00052 \times 0{\cdot}83$$
$$= -0{\cdot}00043.$$

From equation (38), and Table 1, the remaining parameters for the regression lines are

$$\alpha_1 = 5{\cdot}5980 - 2{\cdot}7982 \times 1{\cdot}4417 + 0{\cdot}0005 \times 1{\cdot}1325$$
$$= 1{\cdot}564$$

and
$$\alpha_2 = 5{\cdot}8841 - 2{\cdot}7982 \times 1{\cdot}3017 + 0{\cdot}0005 \times 1{\cdot}3233$$
$$= 2{\cdot}242.$$

Table 1. *Toxicity of derris roots W.213 and W.214 to* Oryzaephilus surinamensis

(p^* and p are shown as percentages in the table, but in formulae are used as proportions)

x	n	r	p^*	p ($c = 17{\cdot}0$)	Empirical probit	Y	nw	t	y	nwx	nwt	nwy
						Root W.213						
2·17	142	142	100·0	100·0	∞	7·6	4·7	0·343	7·94	10·199	1·6121	37·318
2·00	127	126	99·2	99·0	7·33	7·2	9·7	0·392	7·31	19·400	3·8024	70·907
1·68	128	115	89·8	87·7	6·16	6·3	35·0	0·565	6·15	58·800	19·7750	215·250
1·08	126	58	46·0	34·9	4·61	4·6	47·5	1·780	4·61	51·300	84·5500	218·975
							96·9			139·699	109·7395	542·450
						Root W.214						
1·79	125	125	100·0	100·0	∞	7·2	9·5	0·392	7·59	17·005	3·7240	72·105
1·66	117	115	98·3	98·0	7·05	6·9	14·9	0·438	7·03	24·734	6·5262	104·747
1·49	127	114	89·8	87·7	6·16	6·4	31·4	0·539	6·12	46·786	16·9246	192·168
1·17	51	40	78·4	74·0	5·64	5·5	22·9	0·876	5·64	26·793	20·0604	129·156
0·57	132	37	28·0	13·3	3·89	3·8	17·6	4·557	3·89	10·032	80·2032	68·464
							96·3			125·350	127·4384	566·640
						Controls						
—	129	21	16·3	—	—	—	—	—	—	—	—	—

$W.213$: $\bar{x}_1 = 1{\cdot}4417$, $\bar{t}_1 = 1{\cdot}1325$, $\bar{y}_1 = 5{\cdot}5980$.
$W.214$: $\bar{x}_2 = 1{\cdot}3017$, $\bar{t}_2 = 1{\cdot}3233$, $\bar{y}_2 = 5{\cdot}8841$.

	Snwx²	*Snwxt*	*Snwt²*	*Snwxy*	*Snwty*	*Snwy²*
W.213	215·1198	135·6391	163·7154	820·907	551·987	3147·90
	201·4016	158·2095	124·2803	782·040	614·326	3036·66
	13·7182	−22·5704	39·4351	38·867	−62·339	111·24
W.214	178·2746	111·9036	396·4995	779·415	602·854	3454·48
	163·1633	165·8817	168·6453	737·573	749·862	3334·17
	15·1113	−53·9781	227·8542	41·842	−147·008	120·31
Controls		$n_c(1-C)/C = 629{\cdot}8235$			$n_c(c-C)/C = -5{\cdot}312$	
Totals	28·8295	−76·5485	897·1128	80·709	−214·659	231·55

The revised regression equations are therefore

$$Y_1 = 1{\cdot}564 + 2{\cdot}798x, \quad Y_2 = 2{\cdot}242 + 2{\cdot}798x,$$

and the estimate of the response rate amongst controls is

$$\begin{aligned} C &= 0{\cdot}17 - 0{\cdot}0004 \\ &= 16{\cdot}96\,\%. \end{aligned}$$

Evaluation of Y_1 and Y_2 for the appropriate values of x gives a new set of expected probits which are very close to those in the Y column of Table 1, and the revised estimate of C is almost identical with the earlier approximation. Consequently, there is no need to undertake any further cycle of calculations. As a test of heterogeneity, derived from equation (28),

$$\begin{aligned} \chi^2_{[6]} &= S_{yy} - \beta S_{xy} - \frac{\delta C}{1-C} S_{ty} \qquad (43) \\ &= 231{\cdot}55 - 2{\cdot}7982 \times 80{\cdot}709 - 0{\cdot}00052 \times 214{\cdot}659 \\ &= 5{\cdot}60. \end{aligned}$$

There were 10 dose levels tested, including the controls, and four parameters, α_1, α_2, β, C, have been estimated, leaving 6 degrees of freedom. Clearly there is no indication of heterogeneity. Had χ^2 been larger, the possibility that large contributions came from classes with very small expectations would have required consideration, and grouping of classes in the usual manner might then have been needed. The previous analysis (Finney, 1947*a*, § 28), in which some classes were grouped together, gave

$$\chi^2_{[2]} = 1{\cdot}07.$$

Since there is no heterogeneity, the matrix V in equation (42) gives the variances and covariance of β and $\delta C/(1-C)$. In particular

$$V(\beta) = 0{\cdot}04485,$$

whence $$\beta = 2{\cdot}798 \pm 0{\cdot}212,$$

and $$V(C) = 0{\cdot}0014412 \times (0{\cdot}83)^2,$$

whence $$C = 16{\cdot}96 \pm 3{\cdot}15\,\%.$$

The logarithm of the relative potency of the two derris roots, or the logarithm of the ratio of equally effective doses, is

$$\begin{aligned} M &= \mu_2 - \mu_1 \\ &= (\alpha_1 - \alpha_2)/\beta \\ &= \left\{\bar{y}_1 - \bar{y}_2 - \beta(\bar{x}_1 - \bar{x}_2) - \frac{\delta C}{1-C}(\bar{t}_1 - \bar{t}_2)\right\} \Big/ \beta \qquad (44) \\ &= -0{\cdot}242. \end{aligned}$$

Since

$$\begin{aligned} g &= t^2 V(\beta)/\beta^2 \\ &= 0{\cdot}022 \end{aligned}$$

is small (note that t here is the 5 % unit normal deviate, or 1·96, and has no connexion with the auxiliary variate), the variance of M can be found and used according to standard rules. The result is

$$V(M) = \frac{1}{\beta^2}\left[\frac{1}{{}_1Snw} + \frac{1}{{}_2Snw} + (\bar{l}_1 - \bar{l}_2)^2\, V\left(\frac{\delta C}{1-C}\right)\right.$$
$$\left. + 2(M + \bar{x}_1 - \bar{x}_2)(\bar{l}_1 - \bar{l}_2)\,\mathrm{Cov.}\left\{\beta, \frac{\delta C}{1-C}\right\} + (M + \bar{x}_1 - \bar{x}_2)^2\, V(\beta)\right]$$
$$= \left[\frac{1}{96{\cdot}9} + \frac{1}{96{\cdot}3} + (0{\cdot}1908)^2 \times 0{\cdot}001441 + 2 \times 0{\cdot}102 \times 0{\cdot}1908 \times 0{\cdot}003827\right.$$
$$\left. + (0{\cdot}102)^2 \times 0{\cdot}044848\right] \div 7{\cdot}8288$$
$$= (0{\cdot}010320 + 0{\cdot}010384 + 0{\cdot}000052 + 0{\cdot}000149 + 0{\cdot}000467) \div 7{\cdot}8288 = 0{\cdot}00273. \qquad (45)$$

Hence
$$M = -0{\cdot}242 \pm 0{\cdot}052.$$

By taking antilogarithms of M and of M decreased or increased by $1{\cdot}96 \times 0{\cdot}052$, the statement may be made that $W.213$ has a potency 57·3 % of that of $W.214$, and that the fiducial limits of this estimate are 72·4 and 45·3 %. All these results are the same as in the previously published analysis, but the method of calculation given here is simpler and more straightforward. Calculation of the 'exact' fiducial limits, as required when g is large, presents no special difficulties, and formulae given elsewhere (Finney, 1947*a*, formula (4.7)) may be adapted for the purpose.

(ii) *Wadley's problem*

Wadley (1949) propounds and solves an interesting variant of the standard probit problem. In experiments on the control of immature stages of fruit flies, samples of fruit may be treated and the number of flies which survive and develop may be counted. The total number of flies treated, however, can be discovered only by dissection of the fruit and counting of the dead flies, a laborious procedure. An alternative is to take a parallel sample of untreated fruit and to use the number of flies developing from this as an estimate of the numbers exposed to treatment in the other samples. Under certain assumptions, this problem also comes under the general theory of § 3, and the calculations based upon Wadley's solution can be simplified.

If before treatment insects are distributed at random in the fruit, so that the number in the standard size of sample used in the tests is a sample from a Poisson distribution of mean N, the probability of observing s survivors in a sample subjected to a treatment which has a probability P of killing an insect is

$$P(s) = \sum_{n=s}^{\infty} e^{-N}\left\{\frac{N^n}{n!}\binom{n}{s}P^{n-s}Q^s\right\}$$
$$= e^{-NQ}\frac{(NQ)^s}{s!}. \qquad (46)$$

Now equation (9) was given in terms of the probability of r responses, but it could equally well have referred to the probability of s non-responses and the same theory would have followed. Hence equation (46) shows the general theory to be applicable to the present problem, with
$$J = 0, \quad K = N. \qquad (47)$$

It is not essential that an estimate of N from an untreated sample should be available, any

more than in the problem of § 4 (i) an estimate of C from a control batch was essential, but the precision of estimation of the parameters may be much reduced if this information is lacking.

From equation (46) $$\log F = \text{const.} - NQ + s \log NQ,$$

whence, by differentiating twice with respect to the argument of $F(\ )$ and putting $s = NQ$, the expected value of F''/F is found as

$$\frac{F''}{F} = -\frac{1}{NQ}.$$

Hence, from equation (19), $$W = NZ^2/Q. \tag{48}$$

This is equivalent to a weighting coefficient

$$w = Z^2/Q, \tag{49}$$

multiplied by N, the expected number of flies tested at a dose. The weighting coefficient is dependent only on the form of the tolerance distribution and may easily be tabulated. The auxiliary variate

$$t = -Q/Z \tag{50}$$

is introduced, and equations (26) then give

$$S_{xx} + \frac{\delta N}{N} S_{xt} = S_{xy}, \quad S_{xt} + \frac{\delta N}{N} S_{tt} = S_{ty}. \tag{51}$$

The contributions from a control, untreated, sample are very simple: S_{tt} must be increased by N and S_{ty} by $(s_0 - N)$, where s_0 is the number of flies developing from this sample.

The procedure for estimation of N and the parameters of the tolerance distribution then follows the usual plan. First a provisional value of N is guessed for the data, s_0 being a useful guide to this. The empirical proportion killed at any dose is taken as

$$p = 1 - \frac{s}{N}, \tag{52}$$

from which empirical y-values (whether probits or as defined by some other tolerance distribution) are found. A provisional regression line then leads to values of t and y, from which equations (51) are constructed. The solutions of these, with

$$\alpha = \bar{y} - \beta\bar{x} - \frac{\delta N}{N}\bar{t}, \tag{53}$$

give revised estimates of the three parameters, and iteration may proceed for as many cycles as seem needed. The calculations are exactly equivalent to those given by Wadley (1949), but they simplify the method, especially because they invoke familiar ideas of regression and lead to a test of heterogeneity that does not require the calculation of every expected frequency.

On account of sampling variation in the actual numbers of individuals tested at different doses, values of s for some low doses may exceed N, and the empirical proportion surviving according to equation (52) is then greater than 1. The data for these doses will play their proper part in increasing the estimate of N if the calculations are performed exactly as described here, using a q greater than 1 or a negative p in the formation of each working

equivalent deviate. Though the method of calculation recommended in this paper has a close resemblance to multiple regression methods, it is in reality a method of solving the complicated non-linear maximum likelihood equations; the occurrence of a q greater than unity is an indication that the analogy is not perfect, and not a condemnation of the method.

Once again equations (49)–(53) are perfectly general and apply to any tolerance distribution of the type of (1). Chief interest, however, will attach to the normal tolerance distribution. Wadley has given a table of the weighting coefficient for this distribution; Table 2 is an extension of this comparable to the usual tables of probit weighting coefficients. The

Table 2. *The weighting coefficient, Z^2/Q*

Expected probit Y	Z^2/Q	Expected probit Y	Z^2/Q
1·1	0·000 000 04	5·1	0·34242
1·2	0·000 000 09	5·2	0·36344
1·3	0·000 000 2	5·3	0·38069
1·4	0·000 000 4	5·4	0·39359
1·5	0·000 000 8	5·5	0·40173
1·6	0·000 002	5·6	0·40488
1·7	0·000 003	5·7	0·40296
1·8	0·000 006	5·8	0·39612
1·9	0·000 01	5·9	0·38466
2·0	0·000 02	6·0	0·36904
2·1	0·00004	6·1	0·34983
2·2	0·00006	6·2	0·32770
2·3	0·00011	6·3	0·30338
2·4	0·00019	6·4	0·27760
2·5	0·00031	6·5	0·25109
2·6	0·00051	6·6	0·22452
2·7	0·00081	6·7	0·19848
2·8	0·00128	6·8	0·17348
2·9	0·00197	6·9	0·14993
3·0	0·00298	7·0	0·12813
3·1	0·00443	7·1	0·10829
3·2	0·00647	7·2	0·09051
3·3	0·00926	7·3	0·07482
3·4	0·01302	7·4	0·06118
3·5	0·01798	7·5	0·04948
3·6	0·02439	7·6	0·03958
3·7	0·03251	7·7	0·03132
3·8	0·04261	7·8	0·02452
3·9	0·05491	7·9	0·01899
4·0	0·06959	8·0	0·01455
4·1	0·08677	8·1	0·01103
4·2	0·10648	8·2	0·00827
4·3	0·12863	8·3	0·00614
4·4	0·15300	8·4	0·00451
4·5	0·17926	8·5	0·00327
4·6	0·20692	8·6	0·00235
4·7	0·23540	8·7	0·00167
4·8	0·26398	8·8	0·00118
4·9	0·29189	8·9	0·00082
5·0	0·31831	9·0	0·00057

auxiliary variate differs only in sign from that of equation (36) and may therefore be read from Finney's Table II (1947*a*).

Wadley gives a numerical example of his problem, and the same data may be used to illustrate the method of calculation now proposed. The 'dose' is here measured by length of exposure to treatment. Wadley investigated the probit regression on number of days of exposure; this seems to lead to some inconsistency, for the regression equation so derived shows a considerable (about 20 %) death-rate for day zero. The regression may be effectively linear between 1 and 6 days, yet depart from linearity between 0 and 1, but that would correspond to a peculiar form of time-tolerance distribution. As an alternative, use of the logarithm of the number of days as a dose metameter seems worth trying; this, in fact, appears to be in at least as good agreement with the data as Wadley's supposition, and has the merit of internal consistency.

Table 3. *Effect of duration of treatment on development of fruit flies*

x	s	p ($N=1034$)	Empirical probit	Y	Nw	t	y	Nwx	Nwt	Nwy
0·778	4	0·996	7·65	7·19	95	−0·393	7·47	73·910	−37·335	709·65
0·699	32	0·969	6·87	6·90	155	−0·438	6·87	108·345	−67·890	1064·85
0·602	55	0·947	6·62	6·54	249	−0·507	6·61	149·898	−126·243	1645·89
0·477	158	0·867	6·02	6·07	368	−0·632	6·02	175·536	−232·576	2215·36
0·301	396	0·617	5·30	5·42	409	−0·923	5·29	123·109	−377·507	2163·61
0·000	715	0·309	4·50	4·29	131	−2·455	4·52	0·000	−321·605	592·12
Controls	1070	—	—	—	1407	—	—	630·798	−1163·156	8391·48

$\bar{x} = 0{\cdot}4483, \quad \bar{t} = -0{\cdot}8267, \quad \bar{y} = 5{\cdot}9641.$

$SNwx^2$	$SNwxt$	$SNwt^2$	$SNwxy$	$SNwty$	$SNwy^2$
344·2602	−377·068	1393·38	3995·24	−6430·54	50954·28
282·8046	−521·476	961·57	3762·14	−6937·17	50047·57
61·4556	144·408	431·81	233·10	506·63	906·71
		1034·00		36·00	
		1465·81		542·63	

Table 3 shows the second cycle of calculations for these data. The first cycle had given

$$N = 1034, \quad Y = 4{\cdot}295 + 3{\cdot}726x$$

as approximations to the maximum likelihood estimates, x being the logarithm of the number of days. Consequently, for the second cycle equation (52) gave

$$p = 1 - \frac{s}{1034}$$

as the empirical death-rate. The expected probit, Y, was calculated from the regression equation obtained in the first cycle; in view of the large weights of the observations, two places of decimals were taken. Linear interpolation in Table 2 gave values of w for each Y, and these were multiplied by 1034 to give the total weight for each observation. Since N is the same for each line of the table, this multiplication could have been deferred until the end of the calculations; it is sometimes convenient, however, to have records of the weights of each observation, and this arrangement can take account of differences in the sizes of sample

used at different doses by making the multiplier of w the appropriate fraction of N. Working probits, y, were formed from the table of Finney & Stevens (1948), in order to avoid interpolation for Y, but linear interpolation in other tables would be sufficiently exact.

The remainder of Table 3 follows the familiar pattern. To the sum of squares of deviations of t was added N, and to the sum of products of deviations of t and y was added $(s_0 - N)$, or 36·00. Equations (51), for β and δN, then became

$$61{\cdot}4556\beta + 144{\cdot}408\frac{\delta N}{N} = 233{\cdot}10, \quad 144{\cdot}408\beta + 1465{\cdot}81\frac{\delta N}{N} = 542{\cdot}63.$$

The inverse matrix of the coefficients of β and $\delta N/N$ is

$$V = \begin{pmatrix} 0{\cdot}021\,1735 & -0{\cdot}002\,0860 \\ -0{\cdot}002\,0860 & 0{\cdot}000\,8877 \end{pmatrix},$$

whence were derived $$\beta = 3{\cdot}8036, \quad \frac{\delta N}{N} = -0{\cdot}00455.$$

By substitution of the provisional value of N

$$\delta N = -4{\cdot}7,$$

and the revised estimate of N is therefore 1029·3. Again

$$\begin{aligned} \alpha &= \bar{y} - \beta\bar{x} - \frac{\delta N}{N}\bar{t} \\ &= 4{\cdot}255, \end{aligned}$$

so that the revised regression equation is

$$Y = 4{\cdot}255 + 3{\cdot}804x.$$

The iterative process could be continued from these results, but the new column of expected probits is so nearly the same as that in Table 3 that further calculation is not worth while.

The heterogeneity test (equation (28)) here takes the form

$$\begin{aligned} \chi^2_{[4]} &= 906{\cdot}71 - \beta S_{xy} - \frac{\delta N}{N} S_{ty} \\ &= 22{\cdot}55, \end{aligned} \tag{54}$$

since 4 degrees of freedom remain after the estimation of three parameters. This χ^2 is clearly highly significant, but Wadley found an equally great indication of heterogeneity in his analysis using 'days' instead of 'log days' for x; evaluation of the expected frequencies, NQ, corresponding to each s shows that the large χ^2 is not attributable to a large contribution from one class with small expectation and also that the deviations of s from NQ do not follow any systematic pattern. Genuine heterogeneity of the experimental material, as judged by the standard of the Poisson and binomial distributions used in the derivation of the statistical technique, seems the most likely explanation. Non-normality of the distribution of log tolerances would show itself as a curvature of the probit regression line, but departure from a Poisson distribution, in the numbers of flies per fruit—a situation very likely to obtain in reality—might reduce the true weights of observations and so permit erratic deviations from the regression line which would appear to be significant when the Poisson weights were used. Wadley has already suggested such heterogeneity. The theoretical consequences need

further investigation, but the difficulty is likely to be overcome in large part by the use of a heterogeneity factor (Finney, 1947*a*, § 18). This is

$$\chi^2/4 = 5{\cdot}64,$$

and is assigned 4 degrees of freedom.

From the variance matrix, V,

$$V(N) = 5{\cdot}64 \times 0{\cdot}0008877 \times 1034^2$$
$$= 5353,$$

and therefore

$$N = 1029{\cdot}3 \pm 73{\cdot}2,$$

the standard error being based on only 4 degrees of freedom. The estimation of the mean number of flies exposed in each test is not very reliable and does not differ significantly from the figure of 1070 based on the controls alone. An estimate of LD 50, the 'dose', or in this example the time, expected to kill 50 % of the flies can be derived from m, the value of x which makes $Y = 5$. This is

$$m = (5-\alpha)/\beta \tag{55}$$
$$= 0{\cdot}196.$$

After expressing m in terms of $\bar{y}, \beta, \delta N$, by use of equation (53), its variance can be shown to be

$$V(m) = \frac{1}{\beta^2}\left[\frac{1}{SNw} + \bar{l}^2 V\left(\frac{\delta N}{N}\right) - 2(m-\bar{x})\,\bar{l}\,\mathrm{Cov.}\left(\beta, \frac{\delta N}{N}\right) + (m-\bar{x})^2 V(\beta)\right], \tag{56}$$

when there is no heterogeneity, the variance and covariances being taken from the matrix V. Here, the expression on the right-hand side of equation (56) must be multiplied by the heterogeneity factor, giving

$$V(m) = 5{\cdot}64 \times [0{\cdot}0007107 + (0{\cdot}827)^2 \times 0{\cdot}000888 + 0{\cdot}504 \times 0{\cdot}827 \times 0{\cdot}002086$$
$$+ (0{\cdot}252)^2 \times 0{\cdot}021174] \div (3{\cdot}8036)^2$$
$$= 0{\cdot}00138.$$

Hence the standard error of m is $\pm 0{\cdot}037$. Since the heterogeneity factor is based on only 4 degrees of freedom, this standard error must be multiplied by 2·78 in order to give the width of the 5 % fiducial interval. The fiducial limits to m are therefore 0·299 and 0·093. These values correspond to an estimated LD 50 of 1·57 days, with limits at 1·99 and 1·24 days. The legitimacy of using the standard error may be judged from the criterion

$$g = 5{\cdot}64 \times (2{\cdot}78)^2 \times 0{\cdot}02117/(3{\cdot}804)^2$$
$$= 0{\cdot}064.$$

When g is less than 0·1, for most practical purposes the approximate method for obtaining the limits is sufficiently good; application of Fieller's method for calculation of the true fiducial limits (Finney, 1947*a*, formula (4.7)) gives these as 1·94 and 1·17 days.

(iii) *Related problems*

The adjustment of standard probit analyses in order to take account of natural mortality, and the maximum likelihood analysis of Wadley's problem are instances in which the theorem of § 3 is exactly applicable. Other problems can be treated by very closely related methods. The fitting of the Parker-Rhodes equation (Finney, 1947*a*, § 45), for example, requires that equation (8) be replaced by

$$Y = \alpha + \beta x^i, \tag{57}$$

where x is now the absolute dose, not its logarithm, and i is a third parameter requiring estimation. The maximum likelihood equations, which the writer has derived elsewhere, can be rearranged so as to make the calculations formally equivalent to those for a multiple regression on x and an auxiliary variate

$$t = x^i \log x. \tag{58}$$

Quantitative responses whose relationship to dose may be taken as proportional to an integral such as that in equation (3) may also be analysed by a similar procedure; the method follows the lines suggested elsewhere (Finney, 1947*a*, § 47), but again simplifies the scheme of calculation by the introduction of an auxiliary variate

$$t = P/Z, \tag{59}$$

with the aid of which a close analogy with multiple regression is brought about. Previous accounts of both these problems discussed them only for the case in which (1) is a normal distribution, but from the present paper it is clear that the methods can be readily adapted for use with other distributions dependent only upon a parameter of location and a parameter of scale.

5. Summary

Existing methods for the analysis of data relating the dose level of some stimulus to the proportion of individuals showing a characteristic response are in effect methods for estimating the parameters of an underlying tolerance distribution. Usually the distribution is assumed to be normal in respect of some known dose metameter, and the probit method may then be used. Analogous computational procedures can be used for any other tolerance distribution which is completely specified by a parameter of location and a parameter of scale. In the present paper a general method is developed, applicable to any such distribution, which leads to a convenient computational routine when experimental conditions prevent direct observations of proportions responding. Instances of practical importance are the adjustment of data required when a 'natural' response rate is superimposed on that due to the stimulus, and the Wadley problem in which the numbers of individuals exposed to the stimulus must be estimated from a parallel sample instead of counted in the sample tested. When certain simple supplementary tables have been prepared, these and related problems can be dealt with by iterative calculations of the type now familiar in connexion with the standard probit method. The process becomes formally identical with that used for multiple linear regression, except that both the dependent and some independent variates are modified at the end of each cycle of iteration, instead of the unusual patterns of computation previously recommended.

REFERENCES

Finney, D. J. (1944). The application of the probit method to toxicity test data adjusted for mortality in the controls. *Ann. Appl. Biol.* **31**, 68–74.

Finney, D. J. (1947*a*). *Probit Analysis: A Statistical Treatment of the Sigmoid Dose Response Curve.* Cambridge University Press.

Finney, D. J. (1947*b*). The principles of biological assay. *J. Roy. Statist. Soc. Suppl.* **9**, 46–91.

Finney, D. J. & Stevens, W. L. (1948). A table for the calculation of working probits and weights in probit analysis. *Biometrika*, **35**, 191–201.

Horsfall, J. G. (1945). *Fungicides and their Action.* Waltham, Mass., U.S.A.: Chronica Botanica Co.

Wadley, F. M. (1949). Dosage-mortality correlation with number treated estimated from a parallel sample. *Ann. Appl. Biol.* **36**, 196–202.

AN OVERLAP PROBLEM ARISING IN PARTICLE COUNTING

By P. ARMITAGE, B.A.

Medical Research Council Statistical Research Unit,
London School of Hygiene and Tropical Medicine

Introduction

1. The problem with which this paper is concerned has already been briefly discussed by Irwin, Armitage & Davies (1949). In counts of dust particles on a sampling plate, the total number of particles present may be underestimated on account of the overlapping of some of the particles. 'Clumps' of particles formed in this way cannot be analysed under the microscope, and will be counted as a single particle instead of two, three, or more. A similar situation arises in bacterial counting, where the number of organisms present on a plate is obtained by counting the number of colonies which develop from them in a certain length of time. Two or more colonies which overlap may be counted as a single colony, in which case the number of colonies present will be underestimated.

The purpose of the present investigation is to obtain a method of correcting an observed count of this sort, so as to allow for the possibility of undetected overlaps. A considerable gap between theory and practice is probably inevitable, at least in the application to the counting of dust particles, because of the variation in the size and shape of the particles. It is precisely this variation which makes it impossible for an experimenter to distinguish a clump from a single particle.

The simplest mathematical model is that suggested by Irwin *et al.* The particles are regarded as circular laminae of equal diameter δ, and we consider the formation of clumps when N particles fall at random on an area A. The concentration of the particles on the plate may be measured by the quantity

$$\psi = \pi N \delta^2 / 4A$$

(ψ is the ratio of the sum of the areas of the particles to the area of the plate). We shall assume in the discussion of this model that for a given concentration ψ, N is sufficiently large, and δ^2/A sufficiently small, to make it justifiable, on the one hand, to neglect certain 'edge effects' due to the presence of particles near the boundary of the plate, and on the other hand to deal merely with the expected numbers of clumps of various sizes, neglecting completely the sampling variation of these numbers.

It may be possible to avoid any bias due to edge effects by adopting the convention that clumps overlapping one half of the boundary are included in the count, while those overlapping the other half are not included. In any case, for a given concentration ψ, the edge effects assume diminishing importance as N increases and δ^2/A decreases.

In the note previously referred to, Irwin *et al.* have given an approximate formula for the mean clump size. Defining $m = N/C$, where C is the number of clumps on the plate, their formula is (in the present notation)

$$m = \frac{4\psi}{1 - e^{-4\psi}}, \tag{1}$$

$$= 1 + 2\psi + O(\psi^2). \tag{2}$$

The argument by which (1) is obtained involves the assumption that, in each clump, every particle overlaps every other one. This is not necessarily true for clumps of more than two particles, and we shall show that (1) underestimates the true value of m, and is actually valid only as far as the first two terms of the expansion given by (2). In §§ 2–4 of the present paper, an expression for m is obtained which is valid to order ψ^2. For small values of ψ, such as should occur in a well-planned particle counting experiment, (1) gives, for most practical purposes, a sufficiently good approximation to the true value of m for this model.

Two other models are considered in §§ 5 and 6, in which the particles are regarded (*a*) as circular laminae of different sizes, and (*b*) as rectangular laminae of constant proportions and different sizes. Expressions analogous to (1) are obtained in each case, involving the mean and variance of the square root of the area of an individual particle.

Finally, we suggest a formula for the estimation of m in terms of quantities observable in an actual count, which should be applicable for the types of particles usually encountered in practical work.

Circular laminae of equal size

2. As a lemma, we shall need the probability density function (p.d.f.) of the distance r between two points placed at random inside an area A, which is large in comparison with r^2. Garwood (1947) has given the p.d.f. of r, when A is a square, a circle, or a rectangle, of unit area in each case. For a square, for example, the p.d.f. of r is

$$\phi(r) = 2r(\pi - 4r + r^2) \quad \text{for} \quad 0 < r < 1$$

and

$$\phi(r) = 2r(4\sin^{-1} 1/r + 4\sqrt{(r^2-1)} - r^2 - \pi - 2) \quad \text{for} \quad 1 < r < \sqrt{2}.$$

For small r, we see that

$$\phi(r) = 2\pi r + O(r^2). \tag{3}$$

We shall clearly obtain the same limiting result (3) for a unit area A of any shape, provided that the smallest chord of A is large in comparison with r (a provision which must be made, for instance, in the case of the rectangle).

The limiting expression (3) may be obtained directly from simple considerations. For convenience we shall denote the magnitude of the area (not necessarily unity) by A. Given the position of one point, r is less than some value r_0 if the second point falls within a circle of radius r_0, with centre at the first point. Neglecting edge effects, the probability of this event is

$$P(r < r_0) = \pi r_0^2/A.$$

Any complications due to the boundary will affect a proportionate area of order $r_0/A^{\frac{1}{2}}$ (provided that r_0 is small in comparison with the smallest chord of A). Hence

$$P(r < r_0) = (\pi r_0^2/A)\{1 + O(r_0^2/A)^{\frac{1}{2}}\}, \tag{4}$$

as $r_0^2/A \to 0$.

On differentiating (4) with respect to r_0, we have

$$\phi(r) = (2\pi r/A)\{1 + O(r^2/A)^{\frac{1}{2}}\}, \tag{5}$$

which will be seen to be equivalent to (3).

3. Any two particles (which we assume to be circular laminae of diameter δ) will overlap if the distance between their centres is less than δ. Let X and Y be the centres of two overlapping particles, and denote by S and T the circles with centres X and Y respectively, and *radius* δ.

A third particle will form a triplet with the given particles if it overlaps either or both of them. Two types of triplets may be defined:

(*a*) A 'chain' triplet, in which the third particle overlaps only one of the given particles. This will occur if the third particle centre falls in one of the two non-overlapping parts of S or T.

(*b*) A 'complete' triplet, in which the third particle overlaps both the others. This will occur if the third particle centre falls in the area common to S and T.

The area of overlap of S and T is

$$2\delta^2 \cos^{-1}(r/2\delta) - r\sqrt{(\delta^2 - r^2/4)}.$$

The probability that a third particle, falling at random on A, forms a chain triplet with the two given ones is therefore

$$P_1(\delta, r) = (2/A)\{\pi\delta^2 - 2\delta^2 \cos^{-1}(r/2\delta) + r\sqrt{(\delta^2 - r^2/4)}\}, \tag{6}$$

and the probability that it forms a complete triplet is

$$P_2(\delta, r) = (1/A)\{2\delta^2 \cos^{-1}(r/2\delta) - r\sqrt{(\delta^2 - r^2/4)}\}. \tag{7}$$

If N particles fall at random on A, the probability that any two particles, chosen at random out of the N, form part of a chain triplet is (asymptotically, for sufficiently large N)

$$K_1(\delta) \sim N\int_0^\delta \phi(r)\, P_1(\delta, r)\, dr, \tag{8}$$

where $\phi(r)$ is given by (5); and the probability that they form part of a complete triplet is (again for sufficiently large N)

$$K_2(\delta) \sim N\int_0^\delta \phi(r)\, P_2(\delta, r)\, dr. \tag{9}$$

Any inaccuracies in (6) and (7) due to edge effects will be accounted for by the asymptotic nature of (8) and (9). From (8), (5) and (6), for a given value of $\psi = \pi N\delta^2/4A$, and sufficiently small values of δ^2/A,

$$K_1(\delta) = (4N/A)\int_0^\delta (\pi r/A)\{1 + O(r^2/A)^{\frac{1}{2}}\}\{\pi\delta^2 - 2\delta^2 \cos^{-1}(r/2\delta) + r\sqrt{(\delta^2 - r^2/4)}\}\, dr.$$

Using the results

$$\int x \cos^{-1} x\, dx = \tfrac{1}{4}(2x^2 - 1)\cos^{-1} x - \tfrac{1}{4}x\sqrt{(1 - x^2)} + \text{constant},$$

and

$$\int x^2\sqrt{(1 - x^2)}\, dx = \tfrac{1}{8}\{\sin^{-1} x - x\sqrt{(1 - x^2)} + 2x^3\sqrt{(1 - x^2)}\} + \text{constant},$$

we have

$$K_1(\delta) = (4N\pi/A^2)\left\{\pi\delta^4/2 - 8\delta^4\int_0^{\frac{1}{2}} x\cos^{-1} x\, dx + 8\delta^4\int_0^{\frac{1}{2}} x^2\sqrt{(1 - x^2)}\, dx\right\}\{1 + O(\delta^2/A)^{\frac{1}{2}}\}$$

$$= \frac{3\sqrt{3}\,\pi}{2}\frac{N\delta^4}{A^2}\left\{1 + O\left(\frac{\delta^2}{A}\right)^{\frac{1}{2}}\right\}. \tag{10}$$

Similarly, from (9), (5) and (7), we find that

$$K_2(\delta) = \pi\left(\pi - \frac{3\sqrt{3}}{4}\right)\frac{N\delta^4}{A^2}\left\{1 + O\left(\frac{\delta^2}{A}\right)^{\frac{1}{2}}\right\}. \tag{11}$$

Now, suppose that in a count of N particles we find, on the average, P_1 isolated particles, P_2 doublets, P_{31} chain triplets, and P_{32} complete triplets. We may evaluate P_1 easily, for the probability that a particle is isolated is the probability that no other particle centre falls within the circle of radius δ with centre coinciding with the given particle centre. For a given ψ, and values of δ^2/A sufficiently small, this probability is

$$\exp(-\pi N\delta^2/A) = \exp(-4\psi).$$

For large N, therefore,

$$P_1 \sim N\exp(-4\psi) = N\{1-4\psi+8\psi^2+O(\psi^3)\}. \tag{12}$$

There are ${}^NC_2$ pairs of particles, and the proportion of this number which overlap but which do not form parts of triplets is

$$\int_0^\delta \phi(r)\,dr - K_1(\delta) - K_2(\delta) \sim \frac{\pi\delta^2}{A} - \frac{\pi(4\pi+3\sqrt{3})}{4}\frac{N\delta^4}{A^2},$$

for large N. Hence

$$P_2 \sim \frac{N^2}{2}\left\{\frac{\pi\delta^2}{A} - \frac{\pi(4\pi+3\sqrt{3})}{4}\frac{N\delta^4}{A^2}\right\} = 2N\left\{\psi - \frac{(4\pi+3\sqrt{3})}{\pi}\psi^2\right\}. \tag{13}$$

In order to evaluate P_{31} and P_{32}, we remark that out of the ${}^NC_2$ pairs of particles, a proportion $K_1(\delta)$ will overlap and form parts of chain triplets, and a proportion $K_2(\delta)$ will overlap and form parts of complete triplets. Since each chain triplet contains two overlapping pairs, and each complete triplet three overlapping pairs, it follows that

$$P_{31} = {}^NC_2 . K_1(\delta)/2 \quad \text{and} \quad P_{32} = {}^NC_2 . K_2(\delta)/3.$$

From (10) and (11), for large N,

$$P_{31} \sim \frac{3\sqrt{3}\,\pi}{8}\frac{N^3\delta^4}{A^2} = \frac{6\sqrt{3}}{\pi}N\psi^2, \tag{14}$$

and

$$P_{32} \sim \frac{\pi(4\pi-3\sqrt{3})}{24}\frac{N^3\delta^4}{A^2} = \frac{2(4\pi-3\sqrt{3})}{3\pi}N\psi^2. \tag{15}$$

From (14) and (15), the total number of triplets, P_3, is given by

$$P_3 = P_{31}+P_{32} = \frac{4(2\pi+3\sqrt{3})}{3\pi}N\psi^2. \tag{16}$$

We have in (12), (13) and (16) obtained asymptotic expressions for P_1, P_2 and P_3, valid for large N. (We shall henceforth use these formulae without necessarily indicating their asymptotic nature.) No account, however, has yet been taken of clumps of higher order than triplets. Since these large clumps each contain at least two triplets, and since a triplet will form part of a clump of higher order only if at least one other particle centre falls sufficiently near the triplet (which event is easily seen to have a probability of order ψ), it follows that the expected numbers of clumps larger than triplets are of order $P_3\psi$, that is, of order $N\psi^3$.

The number of isolated triplets is therefore, from (16),

$$P_3' = P_3 + N.O(\psi^3) = N\left\{\frac{4(2\pi+3\sqrt{3})}{3\pi}\psi^2 + O(\psi^3)\right\}. \tag{17}$$

It is readily verified from (12), (13) and (17) that

$$P_1+2P_2+3P_3' = N\{1+O(\psi^3)\}, \tag{18}$$

as we should expect. The total number of clumps, C, is given by

$$C = P_1 + P_2 + P_3' + N.O(\psi^3) = N\left\{1 - 2\psi + \frac{2(4\pi - 3\sqrt{3})}{3\pi}\psi^2 + O(\psi^3)\right\}. \tag{19}$$

An interesting point is that, if we were concerned merely to obtain the expression (19) for C, we need not evaluate $K_1(\delta)$. For, assuming on *a priori* grounds that (18) is true, we have, to order ψ^2,

$$\begin{aligned} C &= (N - 2P_2 - 3P_3) + P_2 + P_3 \\ &= N - P_2 - 2P_3 \\ &= N - {}^{N}C_2\left\{\int_0^\delta \phi(r)\,dr - K_1(\delta) - K_2(\delta)\right\} - 2\,{}^{N}C_2\{K_1(\delta)/2 + K_2(\delta)/3\} \\ &= N - {}^{N}C_2\left\{\int_0^\delta \phi(r)\,dr - K_2(\delta)/3\right\}, \end{aligned} \tag{20}$$

a formula not involving $K_1(\delta)$ explicitly. It may easily be verified that (20) does in fact give the correct result (19).

4. We are now able to compare various formulae which are available as approximations to $m = N/C$.

For a finite value of ψ, provided N is sufficiently large and δ^2/A sufficiently small, we have, from (19),

$$\frac{1}{m} = 1 - 2\psi + \frac{2(4\pi - 3\sqrt{3})}{3\pi}\psi^2 + O(\psi^3) \tag{21}$$

$$= 1 - 2\psi + 1{\cdot}564\psi^2 + O(\psi^3). \tag{21a}$$

The equation (1) which is proposed by Irwin *et al.* gives, upon expansion,

$$1/m = 1 - 2\psi + 8\psi^2/3 + O(\psi^3), \tag{22}$$

which agrees with (21) to the first degree in ψ. A comparison of the coefficients of ψ^2 in (21) and (22) shows that, at least for sufficiently small ψ, (21) gives the higher value of m. This result was to be expected, since the proof suggested by Irwin *et al.* takes no account of chain overlaps.

It may seem preferable to obtain an expression for m by the following expansion derived from (21):

$$\begin{aligned} m &= \left\{1 - 2\psi + \frac{2(4\pi - 3\sqrt{3})}{3\pi}\psi^2 + O(\psi^3)\right\}^{-1} \\ &= 1 + 2\psi + \frac{2(2\pi + 3\sqrt{3})}{3\pi}\psi^2 + O(\psi^3) \qquad (23) \\ &= 1 + 2\psi + 2{\cdot}436\psi^2 + O(\psi^3). \qquad (23a) \end{aligned}$$

However, (21) may be seen to give a higher value of m than (23), and since (21) undoubtedly underestimates the true value owing to our neglect of overlaps of higher order than triplets, it is presumably safer to use (21) than (23).

In Table 1 are shown the results of some sampling experiments which were performed by Mr C. N. Davies and the author as a verification of (1), before the present results were obtained. Each experimenter placed 200 points randomly on a 100×100 lattice, by means of random sampling numbers. Circles of various sizes (giving different values of ψ) and centred

upon these random points were then drawn, and the clumps of various sizes were counted. The corresponding values of $m = N/C$ are shown in the third column of Table 1. The standard errors which are also given are estimated by

$$\text{S.E.}(m) = s/\sqrt{C},$$

where s is the estimated standard deviation (with divisor $(C-1)^{\frac{1}{2}}$) of the distribution of clump size.

The theoretical values of m obtained from (1), and from (21) (neglecting in the latter formula terms of order ψ^3), are shown in the fourth and fifth columns of Table 1 respectively. It will be observed that the only experimental value of m which appears to differ significantly from either of the theoretical values is that for experiment (*b*) with $\psi = 0{\cdot}016$. Nevertheless, it may not in fact do so, because for low values of ψ the estimated standard error will be fairly highly correlated with the observed value of m. If, therefore, the value $m = 1{\cdot}005$ is smaller than the true value, its standard error will also be underestimated.

Table 1. *Comparison of experimental and theoretical values of m, for different values of* ψ

$\psi = \frac{\pi N\delta^2}{4A}$	Experimental value of m and standard error	Theoretical values of m	
		From (1)	From (21)
0·016	(*a*) 1·026 ± 0·015 (*b*) 1·005 ± 0·005	1·032	1·032
0·024	1·05 ± 0·02	1·05	1·05
0·035	1·08 ± 0·04	1·07	1·07
0·063	(*a*) 1·15 ± 0·03 (*b*) 1·14 ± 0·03	1·13	1·14
0·098	1·28 ± 0·05	1·21	1·22
0·141	1·38 ± 0·07	1·31	1·34
0·192	1·47 ± 0·08	1·43	1·49
0·251	1·71 ± 0·12	1·59	1·68

The values of ψ occurring in dust-particle counting should be considerably less than 0·2. The last two columns of Table 1 show that the difference between (1) and (21) is, for most practical purposes, immaterial.

Circular laminae of unequal size

5. Suppose that the number of circular particles whose diameter lies between δ and $\delta + d\delta$ is $Nf(\delta)\,d\delta$ $(0 < \delta < \infty)$.

The probability that a given particle S of diameter δ is not overlapped by any particle of diameter between δ' and $\delta' + d\delta'$ is

$$\exp\{-(\pi/4A)(\delta+\delta')^2 Nf(\delta')\,d\delta'\}.$$

The probability that S is not overlapped by any other particle is therefore

$$\exp\left\{-\int_0^\infty (\pi/4A)(\delta+\delta')^2 Nf(\delta')\,d\delta'\right\} = \exp\{-(\pi N/4A)(\delta^2 + 2\delta\nu + \nu_2')\},$$

where ν is the mean, and ν_i' the ith moment about zero, of the distribution of δ.

In the notation of § 3,

$$\begin{aligned} P_1 &= \int_0^\infty Nf(\delta)\exp\{-(\pi N/4A)(\delta^2+2\delta\nu+\nu_2')\}\,d\delta \\ &= \int_0^\infty Nf(\delta)\{1-(\pi N/4A)(\delta^2+2\delta\nu+\nu_2')+\ldots\}\,d\delta \\ &= N\{1-(\pi N/2A)(\nu^2+\nu_2')+R\} \\ &= N\{1-(\pi N/2A)(\nu_2+2\nu^2)+R\}, \end{aligned}$$

where ν_2 is the variance of the distribution of δ, and R is of order $(N\nu^2/A)^2$, if we assume that ν_i' is of order ν^{i+j} (a condition likely to be fulfilled by any distributions arising in practice).

As a first approximation to m, correct to the first order in $(N\nu^2/A)$, we may write

$$\left.\begin{aligned} C &= P_1+P_2 \\ &= P_1+(N-P_1)/2 \\ &= N\{1-(\pi N/4A)(\nu_2+2\nu^2)\} \end{aligned}\right\} \tag{24}$$

and

$$m = 1+(\pi N/4A)(\nu_2+2\nu^2).$$

We may note that the value of m given by (24) is midway between the two values obtained by substituting for δ^2 in (2), ν^2 and ν_2' respectively.

The problem of evaluating P_3, in order to obtain an expression for m comparable with (21), has proved difficult, but we may note that, in the same way that (1) proves a better approximation than (2) in the case of equal circles (since the coefficient of ψ^2 in the expansion of (1) is positive), so we may expect the analogous formula in the case of unequal circles to give a better approximation than (24).

By an argument similar to that of Irwin *et al.*, we obtain the approximation

$$m = \frac{4\psi}{1-e^{-4\psi}}, \tag{25}$$

where

$$\psi = \pi N(\nu_2+2\nu^2)/8A. \tag{26}$$

It is convenient to replace ν and ν_2 by the moments μ and μ_2 of $\sqrt{a}$, the square root of the area of the circular particles. Since

$$\sqrt{a} = \frac{\delta\sqrt{\pi}}{2},$$

we have

$$\frac{\mu_2}{\nu_2} = \frac{\mu^2}{\nu^2} = \frac{\pi}{4},$$

and (26) reduces to

$$\psi = N(\mu_2+2\mu^2)/2A. \tag{27}$$

In the case of equal circles, putting $\mu_2 = 0$ and $\mu^2 = \pi\delta^2/4$, (27) reduces to

$$\psi = \pi N\delta^2/4A,$$

in agreement with the original definition of ψ, as we should expect since (1) and (25) are the same formula.

RECTANGULAR LAMINAE OF UNEQUAL SIZE

6. We shall assume that all the rectangles are of the same proportions, with sides $2l$ and $2kl$ in length, k being a constant. Suppose that the p.d.f. of l is $g(l)$ $(0<l<\infty)$.

A given rectangle L with parameter l will be overlapped by a rectangle L' with parameter l', falling at an inclination β to L, if the centre of L' falls within an area Ω whose boundary is shown as the outer thick line in Fig. 1. The inner thick line in Fig. 1 is the boundary of L, and the dotted lines show the limiting positions of L'.

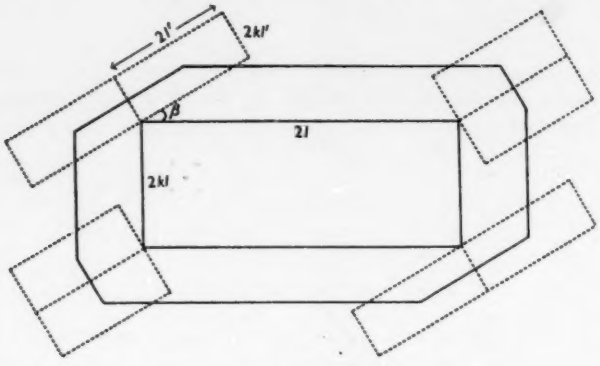

Fig. 1. Diagram showing the admissible area for the centre of a rectangle L' overlapping a rectangle L.

The admissible area Ω is

$$4ll'\{(1+k^2)\sin\beta+2k\cos\beta\}+4k(l^2+l'^2).$$

Assuming that all values of β are equally likely, the probability that L is not overlapped by any other rectangle is therefore

$$\begin{aligned}\exp\left\{-\frac{2}{\pi}\int_0^\infty\int_0^{\frac{1}{2}\pi}(4N/A)\,g(l')\,[\{(1+k^2)\sin\beta+2k\cos\beta\}ll'+k(l^2+l'^2)]\,dl'\,d\beta\right\}\\ =\exp\left\{-\int_0^\infty(4N/\pi A)\,g(l')\,[2(1+k)^2ll'+\pi k(l^2+l'^2)]\,dl'\right\}\\ =\exp\{-(4N/\pi A)\,[2(1+k)^2l\nu+\pi k(l^2+\nu_2')]\},\end{aligned}$$

where ν is the mean, and ν_i' is the ith moment about zero of the distribution of l. Hence

$$\begin{aligned}P_1&=\int_0^\infty Ng(l)\exp\{-(4N/\pi A)\,[2(1+k)^2l\nu+\pi k(l^2+\nu_2')]\}\,dl\\ &=N\{1-(4N/\pi A)\,[2(1+k)^2\nu^2+2\pi k\nu_2']+R\}\\ &=N\{1-(8N/\pi A)\,[\{(1+k)^2+\pi k\}\nu^2+\pi k\nu_2]+R\},\end{aligned}$$

where ν_2 is the variance of the distribution of l, and R is of order $(N\nu^2/A)^2$, assuming that ν_i' is of order ν^{i+1}.

As in § 5, we obtain, as a first approximation to m,

$$m=1+(4N/\pi A)\,[\{(1+k)^2+\pi k\}\nu^2+\pi k\nu_2],$$

and as a better approximation $$m = \frac{4\psi}{1-e^{-4\psi}},$$

where $$\psi = (2N/\pi A)\,[\{(1+k)^2+\pi k\}\,\nu^2+\pi k\nu_2].$$

Replacing ν and ν_2 by the moments μ and μ_2 of $\sqrt{a} = 2l\sqrt{k}$, we have

$$\psi = (N/2A)\left[\mu_2+\left\{1+\frac{(1+k)^2}{\pi k}\right\}\mu^2\right]. \tag{28}$$

We may now compare (27) and (28), since in both formulae ψ is expressed in terms of the moments of the distribution of the square root of the area of the individual particles.

For square particles, putting $k = 1$ in (28), we have

$$\psi = N(\mu_2+2{\cdot}273\mu^2)/2A, \tag{29}$$

a result, as we should expect, not very different from (27).

As $k\to\infty$ or $k\to 0$ in (28), ψ, and therefore m, increases indefinitely. This, also, was to be expected, since needle-shaped particles having areas comparable to those of circular particles, must be of infinite length, and each particle may therefore be expected to cross every other one.

For $k = \frac{1}{5}$ or $k = 5$, we have, from (28),

$$\psi = N(\mu_2+3{\cdot}292\mu^2)/2A. \tag{30}$$

Even for such long rectangles the difference between (27) and (30) is not very serious.

PRACTICAL APPLICATIONS OF THEORY

7. We have, in §§ 5 and 6, proposed that the mean clump size should be estimated by the approximation

$$m = \frac{4\psi}{1-e^{-4\psi}}, \tag{25}$$

and have obtained, in (27) and (28), expressions for ψ appropriate to the cases where the particles are unequal circles, or unequal (but proportionate) rectangles.

The expression for ψ appropriate to any particular practical situation will probably be intermediate between (27) and (30) (the latter formula corresponding to the case of rectangles with $k = \frac{1}{5}$ or $k = 5$, which are much more needle-like than the shapes which are likely to arise in practice). The experimental worker should have some idea of the shapes which are likely to occur. In the absence of such knowledge, perhaps a suitable expression for ψ would be

$$\psi = N(\mu_2+2{\cdot}5\mu^2)/2A = NK/2A, \quad \text{say.} \tag{31}$$

Using (31), then, and assuming that μ and μ_2 are known, and that C and A are observed from the count, we may write $N = mC$, and, from (25),

$$m = \frac{2KmC/A}{1-e^{-2KmC/A}}.$$

Hence $$1-e^{-2KmC/A} = 2KC/A$$

and $$m = \frac{A}{2KC}\log_e\left\{\frac{1}{1-2KC/A}\right\}, \tag{32}$$

which may be used as an equation for the estimation of m.

We have assumed, in obtaining (32), that μ and μ_2 are known, or at least may be estimated with sufficient accuracy. We may, for instance, be able to use results obtained from a previous count on the same type of particles, in which the concentration was so low that overlapping did not occur. If we have to estimate μ and μ_2 from the count itself, as the mean and variance of the square root of the area of a clump, ψ will be overestimated by a factor

$$1+O(\psi).$$

Since (25) is in any case valid only to $O(\psi)$, we may disregard this source of error.

This problem was suggested by Mr C. N. Davies of the Medical Research Council Group for Research in Industrial Physiology, and I am indebted to Mr Davies, to my colleague Dr J. O. Irwin, and to Dr H. O. Lancaster of Sydney for some helpful discussions on the mathematical treatment. I should also like to thank Mrs M. G. Young for preparing the diagram.

REFERENCES

Garwood, F. (1947). The variance of the overlap of geometrical figures with reference to a bombing problem. *Biometrika*, **34**, 1.

Irwin, J. O., Armitage, P. & Davies, C. N. (1949). The overlapping of dust particles on a sampling plate. *Nature, Lond.*, **163**, 809.

TABLES OF AUTOREGRESSIVE SERIES

By M. G. KENDALL

In my *Contributions to the Study of Oscillatory Time-Series* (1946) I gave in full four series which had been calculated from the formula

$$u_{t+2}+au_{t+1}+bu_t=\epsilon_{t+2}, \tag{1}$$

with certain values of a and b and a rectangular random element ϵ. These series have repeatedly been used by other workers to exemplify or to verify theoretical results; see, for example, Yule (1945), Bartlett (1946), Quenouille (1947) and Orcutt (1948). It may, therefore, be useful to publish further series subsequently prepared for some studies which are not yet ready to appear.

I preserve the numbering of the four series already published and of some additional series referred to, but not given, in a later paper (Kendall, 1949). The full set of series, including those now published, is as follows:

No. of series	No. of terms	Nature of stochastic element	Generating equation
1	480	Rectangular	$u_{t+2}-1{\cdot}1u_{t+1}+0{\cdot}5u_t=\epsilon_{t+2}$
2	240	"	$u_{t+2}-1{\cdot}2u_{t+1}+0{\cdot}4u_t=\epsilon_{t+2}$
3	240	"	$u_{t+2}-1{\cdot}1u_{t+1}+0{\cdot}8u_t=\epsilon_{t+2}$
4	240	"	$u_{t+2}+1{\cdot}0u_{t+1}+0{\cdot}5u_t=\epsilon_{t+2}$
5*a*	400	"	As series 1
5*b*	400	"	"
5*c*	400	"	"
5*d*	400	"	"
5 (*in toto*)	1600	"	"
6*a*	400	"	As series 3
6*b*	400	"	"
6*c*	400	"	"
6*d*	400	"	"
6 (*in toto*)	1600	"	"
7	500	Normal	$u_{t+1}-0{\cdot}9u_t=\epsilon_{t+1}$
8	500	Series 7	As series 1
9	500	Normal	$u_{t+1}-0{\cdot}7u_t=\epsilon_{t+1}$
10	500	Series 9	As series 1
11	500	Normal	$u_{t+1}-0{\cdot}5u_t=\epsilon_{t+1}$
12	500	Series 11	As series 1
13	500	Normal	$u_{t+1}-0{\cdot}3u_t=\epsilon_{t+1}$
14	500	Series 13	As series 1
15	500	Normal	$u_{t+1}-0{\cdot}1u_t=\epsilon_{t+1}$
16	500	Series 15	As series 1

The stochastic element for series 1–6 inclusive was obtained by taking two-figure numbers from the *Tables of Random Sampling Numbers* by Babington Smith and myself, ignoring 00 and reducing to zero mean by subtracting 50. The 'normal' element for the odd-numbered series 7–15 was obtained from the tables by Mahalanobis and others (1934), which are to three decimal places and were derived by converting to normal deviates the *Tables of Random Sampling Numbers* by L. H. C. Tippett. (When the work was done H. Wold's (1948)

conversion of the Kendall-Babington Smith numbers was not available.) The values of these series were themselves used as the stochastic elements in constructing the even-numbered series 8–16, so that these latter have the same structural constants as series 1, but are based on a stochastic element which is itself autocorrelated in a Markoff chain.

If η obeys the relation

$$\eta_{t+1} + k\eta_t = \epsilon_{t+1}, \tag{2}$$

then

$$u_{t+2} + au_{t+1} + bu_t = \eta_{t+2}$$

is equivalent to

$$u_{t+3} + (a+k)\,u_{t+2} + (b+ak)\,u_{t+1} + bku_t = \epsilon_{t+3}, \tag{3}$$

that is to say, to a third-order autoregressive series with a random stochastic element. Evidently more complicated series may be built up, if desired, from those here given. Conversely, any linear autoregressive series can be regarded as built up from those of the first or second order by an iterative process in which the stochastic element at each stage is the value of the series generated by the previous stage.

In addition to the series themselves I give, in Tables 19–22, the serial covariances (or rather, the serial first-product-moments-about-zero) and the serial correlations for all series except the odd-numbered series from 7 to 15, which have not been calculated. A few points require explanation:

(*a*) *'Starting up' the series.* Series 1–4, the subseries of series 5 and 6, and the odd-numbered series 7–15 were started up by taking u_{-1} and u_0 to be random numbers of the same character as the stochastic element and beginning the series with u_1. The subseries of series 5 and 6 were started up separately. For the even-numbered series 8–16 u_{-1} and u_0 were taken to be the first two terms of the corresponding odd-numbered series, two extra terms of the latter, numbered 501 and 502, having been computed so as to give 500 terms of the derived series.

(*b*) *Product-moments.* The theoretical mean of all the series is zero (and the actual means for series of these lengths are very close to zero); and there are both practical and theoretical reasons for calculating product-moments about zero rather than the actual means. I denote the first-product-sum-about-zero $\sum_t u_t u_{t+k}$ by c_k and give the values of c_k for various ranges of k in Tables 19–22.

For series 1–4 and 7–16 the sum c_k is taken simply over the values of the series, i.e. is the sum of $n-k$ terms beginning with $u_1 u_{1+k}$ and ending with $u_{n-k} u_n$. To simplify the arithmetic for the subseries 5*a*–6*d* the sum is taken over 400 terms, the extra k being obtained by a circular definition; that is, c_k is the sum of terms beginning with $u_1 u_{1+k}$, proceeding through $u_{500-k} u_{500}$ to $u_{501-k} u_1$, $u_{502-k} u_2$, etc., and ending with $u_{500} u_k$. For the full series 5 and 6, c_k was taken as the sum of the four corresponding values for the constituent series.

(*c*) *Serial correlations.* As an estimate of the variance, I used $c_0 = \sum_1^n u_t^2$. The estimate of the kth serial correlation r_k is then

$$r_k = \frac{c_k}{c_0} \frac{n}{n-k}, \tag{4}$$

but where, as in series 5 and 6 and the subseries thereof, the summation of c_k is over n terms, the factor $n/(n-k)$ is replaced by unity.

The estimated serials for series 1–4 published in my brochure were based on the formula

$$r_k = \frac{\operatorname{cov}(u_t, u_{t+k})}{\sqrt{\{\operatorname{var} u_t \operatorname{var} u_{t+k}\}}}, \tag{5}$$

and hence may differ slightly from the values of Tables 19 and 20. The differences are negligible for most purposes, at least for series of these lengths, but for short series (4) is better than (5), which may introduce a substantial bias.

I am very grateful to the National Institute of Social and Economic Research for allowing me to call on the help of Miss Joan Ayling in the heavy computing involved in the preparation of these tables. Miss Ayling did most of the arithmetical work in constructing series 5–16 and computing the serial coefficients, and it is a pleasure to record my indebtedness to her.

REFERENCES

BARTLETT, M. S. (1946). *J.R. Statist. Soc. Suppl.* **8**, 27.
KENDALL, M. G. (1949). *Proceedings of the International Statistical Institute at Washington*, 1947 (in the Press).
MAHALANOBIS, P. C. and others (1934). *Sankhyā*, **1**, 289.
ORCUTT, G. H. (1948). *J.R. Statist. Soc.* Series B, **10**, 1.
QUENOUILLE, M. H. (1947). *J.R. Statist. Soc.* **110**, 123.
WOLD, H. (1948). *Tracts for Computers, No. XXV*. Cambridge University Press.
YULE, G. U. (1945). *J.R. Statist. Soc.* **108**, 208.

Table 1. *Series 5a*

No. of term	0+	50+	100+	150+	200+	250+	300+	350+
1	57	10	9	35	−27	−67	−46	−56
2	69	26	6	22	−53	−50	−59	−31
3	65	16	34	54	−67	24	−64	−22
4	80	2	64	53	−86	32	−88	5
5	53	25	18	14	−83	−14	−95	57
6	−6	60	−31	−32	−67	−12	−29	109
7	−55	49	−77	−29	−43	−23	−17	82
8	−72	3	−65	3	−3	−53	−9	57
9	−71	21	−73	65	66	−38	−40	−1
10	−86	−5	−69	70	30	7	−57	−44
11	−68	−6	1	38	22	61	−8	−29
12	−51	32	76	56	10	75	69	−25
13	−48	29	91	65	−16	47	37	−35
14	−59	53	74	64	6	6	46	12
15	−81	58	35	39	17	−59	9	60
16	−99	26	−11	31	59	−74	14	65
17	−31	34	−24	57	27	−83	−5	40
18	5	−19	−1	50	−9	−84	−34	−13
19	66	−59	−15	11	−32	−2	−50	−83
20	51	−91	−59	22	−74	−9	−58	−58
21	65	−65	−17	48	−71	22	−80	−19
22	23	−9	−11	79	−64	−20	−92	−5
23	10	60	−44	107	−64	−65	−89	−12
24	44	22	−90	94	−41	−14	−78	−48
25	82	−16	−47	15	20	29	−66	−32
26	95	−27	−15	10	65	81	−58	−16
27	59	20	27	33	101	60	−12	−48
28	25	40	−1	48	124	−18	8	−79
29	23	32	−58	11	62	−50	−33	−97
30	57	14	−51	−25	−26	−54	−48	−56
31	76	26	−29	−49	−37	−65	−37	−52
32	75	−13	−43	−5	−55	−53	−31	13
33	33	16	−14	31	−22	−46	0	3
34	30	−2	−7	35	−2	−19	14	−33
35	−20	35	28	8	17	−30	2	3
36	−34	89	27	−40	16	13	44	60
37	−65	126	−28	−74	−19	47	22	42
38	−21	76	−11	−35	4	74	−21	50
39	39	68	−2	−47	60	19	−81	18
40	31	44	−28	−43	38	−31	−108	−48
41	55	21	−1	−70	−35	−44	−92	−29
42	42	−21	32	−55	−95	−50	−38	−43
43	21	−66	61	−49	−72	−60	34	−68
44	−5	−36	42	20	−50	−87	43	−55
45	−5	25	19	94	−3	−83	18	12
46	−36	32	9	92	48	−71	−23	52
47	−4	13	29	51	45	12	−68	10
48	−32	−4	5	41	−18	70	−97	−55
49	−19	21	−48	55	−72	25	−90	−37
50	−24	13	−24	14	−90	−25	−58	−59

Thus, the 242nd term lies in the column headed '200+' and the row 42, i.e, is −95.

Table 2. *Series 5b*

No. of term	0+	50+	100+	150+	200+	250+	300+	350+
1	−34	−45	−28	111	−16	−29	−23	16
2	10	−8	−34	37	34	−30	−24	74
3	67	−15	−11	30	1	−61	6	109
4	43	−48	19	−9	−13	−36	−25	35
5	25	−2	49	−74	−11	−48	1	−39
6	33	−10	51	−61	−28	−46	−35	−86
7	69	−31	−7	7	11	−31	−25	−120
8	78	−11	−6	26	22	−40	−44	−80
9	96	29	18	11	68	−71	−2	−17
10	49	35	8	23	51	−50	−25	−18
11	47	28	−48	−26	−3	−54	−64	15
12	43	17	−90	−29	−31	−33	−47	25
13	12	−11	−63	−23	−47	−28	24	28
14	−37	−62	−46	−10	−71	−58	67	50
15	−73	−75	13	1	−67	−19	61	4
16	−45	−16	67	−16	−49	51	−11	−52
17	−40	10	95	−32	−16	81	−11	−85
18	−30	13	43	7	9	107	−21	−89
19	−20	13	−33	67	−6	116	24	−101
20	−23	23	−63	67	−49	113	4	81
21	−27	50	−13	19	−71	40	22	5
22	−51	33	−9	−36	−103	2	53	88
23	7	−10	22	−23	−65	−53	82	97
24	77	−48	32	−53	12	−81	76	19
25	56	−96	39	−57	24	−75	62	−32
26	72	−92	9	−35	20	−11	−7	−13
27	86	−48	−10	−15	10	−4	−44	−28
28	72	15	1	48	−20	5	−23	−24
29	75	81	26	88	−1	28	−43	−34
30	15	119	−7	117	2	57	−2	1
31	−17	134	−69	113	−13	25	57	−6
32	−27	67	−56	96	−4	−48	88	15
33	−3	26	−13	50	−31	−33	89	16
34	46	−54	56	34	−28	18	63	−23
35	91	−89	104	4	−29	−6	11	−69
36	125	−23	95	26	−43	−20	−49	−58
37	81	−18	55	6	16	−32	−29	−10
38	16	18	59	−6	73	−69	−29	6
39	−6	78	25	31	105	−77	16	6
40	24	52	−22	21	37	−18	42	−36
41	64	19	13	−34	−6	28	45	−15
42	39	−13	10	−77	−37	76	18	−8
43	−23	−28	−23	−20	−20	119	−13	2
44	−65	−6	−37	−9	24	62	13	−1
45	−38	45	−59	32	1	1	−9	−2
46	−56	70	−79	77	−20	−6	13	17
47	−87	11	−43	32	−61	24	42	−25
48	−47	12	27	−42	−33	−9	1	−79
49	−47	40	92	−59	19	−39	−44	−85
50	−29	22	110	−41	−4	−81	−24	−36

Table 3. *Series 5c*

No. of term	0+	50+	100+	150+	200+	250+	300+	350+
1	−26	−55	−20	40	−13	14	27	−8
2	−20	−43	−21	100	1	−5	42	−53
3	−6	−37	13	45	32	−18	0	−41
4	42	−15	45	−45	23	−35	2	1
5	32	−4	−3	−96	17	−50	−8	19
6	−13	1	−55	−36	53	−30	−18	52
7	12	−35	−79	−25	19	10	−27	12
8	49	−67	−26	−49	−23	40	−31	−36
9	14	−47	−33	−25	−30	43	−20	−40
10	−53	16	−29	27	−3	9	2	10
11	−99	67	1	−5	9	14	57	−4
12	−57	55	57	−36	−17	14	95	0
13	−12	−22	102	−10	−37	20	95	41
14	12	−22	65	25	−78	−1	53	92
15	22	−27	30	18	−59	−53	−26	41
16	66	−15	−3	−30	17	−104	−32	−37
17	25	−25	11	−60	34	−68	−15	−98
18	−33	−21	1	−2	72	23	42	−53
19	−52	−37	−13	13	109	65	85	22
20	−30	−7	−51	10	83	41	82	80
21	−17	41	−72	−13	24	21	22	105
22	−36	3	−29	−20	−33	44	−30	43
23	17	−39	−30	−52	−24	9	−16	−28
24	36	−78	30	−25	−48	−18	39	−98
25	64	−53	33	−3	−87	−32	87	−134
26	28	−33	−25	51	−73	−19	114	−82
27	14	−45	−92	42	−71	−19	57	−60
28	−15	−67	−65	33	−34	−24	−18	−1
29	−44	−96	8	41	−18	−18	−67	74
30	−30	−115	32	39	34	20	−53	36
31	−54	−79	45	25	75	28	−18	30
32	−72	−14	13	53	106	−12	−33	29
33	−66	11	−23	28	97	−14	−61	18
34	−73	−11	−3	37	21	−11	−85	19
35	−73	16	17	6	3	−22	−21	−14
36	5	−24	21	−17	36	−22	7	−22
37	3	6	23	−35	8	−58	41	1
38	38	−13	−7	−76	9	−35	13	16
39	53	−22	−49	−73	1	3	25	−18
40	47	17	−93	−27	−39	8	2	−70
41	22	−16	−48	−4	−62	30	22	−63
42	−9	−69	−28	29	−15	47	57	−2
43	−62	−31	−47	−10	2	79	75	29
44	−50	41	−60	−4	−22	24	63	37
45	10	93	−47	−11	−68	−34	−3	−17
46	55	84	25	−48	−26	0	−83	−79
47	64	64	30	−67	−41	65	−135	−70
48	92	0	12	−94	−44	97	−93	−43
49	95	−81	−34	−116	9	30	−11	1
50	29	−92	−16	−61	13	24	1	52

Table 4. *Series 5d*

No. of term	0+	50+	100+	150+	200+	250+	300+	350+
1	8	68	−52	63	18	−12	69	−9
2	24	117	−79	39	11	−28	22	27
3	21	81	−49	23	18	−64	−52	−5
4	38	−17	−2	−8	41	−42	−52	−61
5	45	−72	60	−33	77	−24	17	−52
6	−9	−29	49	11	51	18	2	20
7	−42	−42	22	47	29	73	−26	66
8	−5	−55	18	9	16	67	−11	54
9	−25	8	51	−60	−21	38	−3	6
10	21	68	20	−96	9	−18	−5	29
11	36	59	−12	−36	17	−57	−23	11
12	0	5	−10	30	−19	−27	6	−11
13	−47	23	−30	76	−56	−11	18	17
14	−52	28	−62	38	−43	44	−19	5
15	−49	10	−71	10	−45	87	−11	18
16	5	39	−90	−52	3	121	44	−12
17	54	60	−63	−111	29	79	54	−26
18	74	14	12	−97	37	24	72	−8
19	56	19	45	−88	0	−16	99	−15
20	63	−26	78	−32	1	−9	48	1
21	−3	−4	33	39	8	−40	39	−11
22	−42	−32	42	33	−23	−52	−18	35
23	−75	2	24	−32	12	−52	−30	70
24	−89	51	−25	−56	11	−64	20	77
25	−86	23	−27	−28	17	−60	86	22
26	−27	32	19	−28	−3	−78	116	−9
27	16	67	22	−51	−16	−30	109	−58
28	56	34	27	−77	−10	0	19	−32
29	77	−11	7	−19	−9	56	−9	23
30	19	−45	−16	16	−13	64	−8	4
31	−43	−34	−13	−11	−36	−1	−18	−24
32	−22	−60	−39	−24	−39	−46	4	14
33	28	−63	9	22	−42	−79	−7	64
34	86	3	65	20	−24	−102	3	59
35	117	1	114	29	−20	−84	−39	−9
36	108	9	68	53	31	−29	−64	−36
37	103	50	48	67	17	57	−5	−60
38	53	7	4	47	−39	94	33	−58
39	−3	−23	12	−4	−3	104	58	−62
40	−29	−66	−14	−68	23	115	89	5
41	−57	−90	−55	−45	21	41	37	11
42	−92	−20	−25	−53	−12	−33	20	−3
43	−60	3	−2	−35	−44	−67	3	10
44	4	58	−27	11	−12	−88	2	−25
45	14	33	−29	34	24	−76	−45	−22
46	53	−36	−55	41	24	−42	−73	−46
47	11	−73	−83	18	36	−17	−96	−13
48	33	−96	−40	11	25	−8	−70	−1
49	−18	−95	29	14	50	49	−16	−20
50	12	−50	62	6	4	99	18	−70

Table 5. *Series 6a*

No. of term	0+	50+	100+	150+	200+	250+	300+	350+
1	46	−36	−16	−5	147	98	48	−21
2	−16	55	−71	54	121	42	16	−14
3	−37	119	−96	80	−13	−36	−10	−22
4	−19	78	−90	15	−83	−105	−36	0
5	17	38	11	−42	−97	−134	−26	51
6	52	−14	129	−44	−78	−42	−45	29
7	51	−35	93	23	−13	70	−56	−42
8	47	8	46	104	26	108	−28	−64
9	−8	71	−28	70	−5	62	58	11
10	−11	40	−79	0	−73	17	87	95
11	−50	−50	−52	−16	−93	5	84	48
12	−25	−112	−21	−30	−12	5	38	−36
13	30	−65	55	−36	78	−15	17	−101
14	83	35	89	19	61	−62	−13	−79
15	24	69	47	86	−42	−14	9	11
16	−73	62	−37	39	−103	21	2	124
17	−122	−2	−114	−51	−86	67	−53	170
18	−49	−54	−93	−47	−24	93	−100	105
19	47	−76	18	−21	30	67	−66	−37
20	44	−36	75	−2	29	−43	32	−111
21	−22	54	32	48	−37	−55	40	−55
22	−32	127	2	82	−20	−38	11	46
23	9	106	−38	8	−13	−40	−27	128
24	6	−29	−55	−50	−10	−37	−4	129
25	−18	−141	−39	−33	−26	24	18	12
26	−48	−117	−30	29	21	84	−9	−71
27	−72	−18	−33	14	67	92	−57	−107
28	−14	55	−32	−36	19	30	−98	−16
29	26	96	−4	−85	11	−89	−93	100
30	87	60	17	−63	44	−132	21	122
31	52	−3	18	40	0	−69	143	8
32	−24	−8	42	137	−71	30	142	−118
33	−20	−48	32	151	−47	76	53	−179
34	−24	−90	−42	87	−6	98	−28	−61
35	−12	−100	−111	−39	23	25	−105	37
36	43	−8	−113	−118	−15	−55	−70	85
37	101	38	−19	−121	−65	−62	−5	76
38	37	36	115	−10	−19	−28	94	26
39	−44	7	142	91	68	−9	65	13
40	−34	−16	64	100	101	−30	36	1
41	−15	11	−44	10	10	−75	15	−36
42	10	17	−148	−58	−24	−59	36	−15
43	53	−20	−100	−44	−49	−49	55	3
44	90	−71	26	18	5	−21	10	23
45	86	−40	118	31	22	44	−28	16
46	39	12	98	−27	−19	28	−39	−26
47	−7	24	−12	−84	9	−40	−16	5
48	−79	2	−71	−120	11	−46	19	−8
49	−126	22	−62	−22	34	−2	6	13
50	−120	22	−58	102	70	54	−8	23

Table 6. *Series 6b*

No. of term	0+	50+	100+	150+	200+	250+	300+	350+
1	34	−51	−61	69	−4	19	−1	64
2	15	−84	−64	75	−63	−35	35	30
3	25	−94	−16	−21	−52	−73	83	−22
4	27	−49	49	−106	38	−88	33	−26
5	55	45	105	−65	47	−23	−72	−10
6	60	126	34	−20	17	18	−149	36
7	49	101	−80	16	28	23	−92	59
8	31	−35	−70	37	11	30	−24	38
9	−26	−165	33	22	−49	−29	81	39
10	−27	−124	75	−2	−26	−16	85	−20
11	−38	21	23	−11	−38	6	19	−6
12	24	168	−24	−26	8	18	−72	41
13	18	187	−50	−26	40	59	−90	49
14	−12	69	−5	−9	54	89	−49	47
15	−12	−53	16	54	67	78	33	46
16	−42	−104	21	78	−5	6	108	49
17	−34	−108	7	67	−90	−40	66	50
18	−12	−32	19	−19	−141	−45	−2	−29
19	39	12	−15	−38	−126	−7	−12	−95
20	83	−5	−66	−42	−9	44	24	−78
21	63	−8	−62	27	42	99	42	−33
22	24	−49	34	98	63	70	−10	−8
23	−58	−71	45	53	−12	20	−89	8
24	−72	−29	52	−53	−86	−7	−133	59
25	−24	6	−28	−71	−98	−15	−37	42
26	−6	−14	−66	−21	−6	−41	38	30
27	−4	4	−66	78	42	−78	58	36
28	−48	44	−28	145	18	−37	−13	−8
29	−74	94	13	117	18	28	−27	−44
30	2	88	59	−7	−33	30	11	−14
31	103	69	63	−108	−21	52	53	9
32	99	53	−9	−70	−1	44	92	42
33	−20	47	−71	52	−26	5	38	6
34	−134	14	−88	141	−46	11	−60	−61
35	−135	27	−37	102	−12	31	−112	−50
36	−65	12	64	22	34	74	−59	−13
37	23	−54	76	−68	22	86	63	12
38	114	−44	11	−76	15	81	160	34
39	62	33	−30	−47	−28	65	77	51
40	15	91	−79	49	−29	1	−89	61
41	−63	74	−107	127	−22	−92	−152	18
42	−119	22	−34	116	25	−59	−75	−48
43	−121	−82	1	64	48	−20	87	−86
44	−42	−61	12	−32	12	65	153	−44
45	80	16	12	−117	−61	77	75	−25
46	94	34	−16	−117	−125	−14	−32	−5
47	27	63	−61	1	−92	−126	−46	53
48	−25	48	−22	68	19	−142	−3	59
49	−5	−39	33	119	76	−57	−5	−9
50	−25	−52	43	91	38	2	15	−36

Table 7. *Series 6c*

No. of term	0+	50+	100+	150+	200+	250+	300+	350+
1	−19	−167	3	−95	10	−67	49	−21
2	−53	−132	−65	−111	45	42	53	−26
3	−9	−20	−31	−29	51	77	59	−59
4	8	91	42	51	38	56	−18	−76
5	−29	103	43	66	−28	26	−26	−68
6	1	45	17	42	−91	8	−64	−23
7	5	−26	−53	−10	−52	−58	−14	−2
8	35	−110	−108	−21	50	−27	77	−12
9	34	−112	−77	20	114	47	51	−23
10	8	−54	20	72	99	91	29	1
11	30	−16	45	52	59	82	−55	67
12	69	9	9	−20	1	2	−114	85
13	85	29	−59	−38	−65	−94	−109	81
14	−12	32	−27	−36	−72	−96	−4	26
15	−89	4	3	28	−64	16	131	1
16	−131	−37	42	102	−32	54	122	−31
17	−57	−79	90	109	60	24	5	7
18	47	−41	35	36	85	−30	−60	−8
19	106	−7	4	−60	64	−48	−82	−33
20	47	−4	19	−108	−40	−7	−62	−24
21	−65	−37	52	−29	−106	12	41	6
22	−97	−54	20	93	−122	19	54	62
23	−89	−2	−16	109	−49	45	3	93
24	−69	68	−48	44	23	18	−5	98
25	7	71	−60	6	97	21	−2	4
26	56	27	−17	−4	124	47	8	−116
27	43	−51	55	−40	16	−3	50	−136
29	34	−29	24	−4	−105	−81	35	−58
29	−33	55	−9	35	−157	−126	33	62
30	−25	108	−38	59	−55	−114	−41	137
31	−4	84	−43	65	111	−11	−26	55
32	36	39	−27	−9	214	98	−13	0
33	63	−24	44	−29	173	83	51	−26
34	68	−87	104	16	43	46	82	5
35	24	−43	39	75	−49	−14	29	49
36	−42	39	−37	64	−94	−32	7	51
37	−76	45	−37	−27	−65	−42	−14	46
38	−25	39	−27	−104	−38	28	−7	−19
39	54	53	−46	−77	−7	67	28	−60
40	128	−18	−68	−43	32	78	73	−19
41	124	−32	3	−31	25	−13	106	27
42	−3	−8	66	−25	−26	−108	89	87
43	−122	20	103	21	−45	−69	43	71
44	−166	7	44	23	−37	31	19	0
45	−120	9	−80	0	11	105	−50	−58
46	−22	24	−101	−5	69	125	−94	−108
47	120	23	−63	−20	60	10	−36	−86
48	176	21	15	25	−30	−68	−2	−49
49	48	38	49	19	−94	−45	11	−2
50	−98	28	−6	−16	−100	−23	−9	−1

Table 8. *Series 6d*

No. of term	0+	50+	100+	150+	200+	250+	300+	350+
1	−57	88	118	−78	−16	−57	7	−35
2	−54	88	106	25	−26	−11	42	−58
3	−41	73	50	68	−32	−2	64	−85
4	26	22	16	30	−2	23	3	−44
5	27	−26	26	−33	12	51	−20	12
6	1	−2	39	−60	27	73	−24	46
7	12	−13	44	−19	38	−8	−19	35
8	13	−47	−2	66	32	−113	46	29
9	−3	−30	1	126	52	−159	87	−27
10	9	−20	−36	104	29	−36	79	−81
11	1	−26	−35	43	15	40	−22	−20
12	−2	−12	−10	−18	21	42	−101	17
13	45	3	46	−98	46	−26	−90	6
14	77	42	10	−73	54	−61	−52	22
15	53	43	−21	5	−4	−25	59	49
16	17	48	−42	68	−15	−20	81	19
17	25	32	−52	55	−52	23	29	−4
18	18	37	−7	−18	−14	67	−16	−33
19	−26	40	41	−26	44	9	−5	8
20	−62	8	38	−44	27	1	−36	−4
21	−84	−67	19	−14	−41	6	8	8
22	−11	−52	−28	68	−111	1	8	48
23	70	−20	−36	42	−92	8	6	65
24	106	−30	−45	17	−3	26	−3	80
25	94	−22	−62	7	60	−5	−3	81
26	63	48	−24	−17	114	19	−6	−14
27	4	101	18	−23	85	62	−40	−78
28	−89	46	22	−52	1	88	−72	−77
29	102	−53	54	−52	−103	81	−46	−65
30	−18	−76	32	−31	−162	8	−20	−51
31	96	6	6	−21	−110	−61	−13	18
32	96	38	−44	−44	31	−78	47	99
33	37	21	−53	−75	118	−86	16	140
34	−28	−21	−33	−5	111	−40	−20	92
35	−61	6	−8	27	−13	−8	−25	−11
36	−85	37	14	−8	−153	54	−11	−74
37	−4	−13	30	−43	−171	41	41	−63
38	74	−73	26	−33	−45	−28	62	30
39	96	−99	−37	13	98	−24	78	58
40	71	−94	−46	88	134	−20	80	41
41	14	−22	−58	59	108	−11	49	43
42	−7	69	−70	−46	10	46	37	61
43	25	87	19	−53	−60	59	47	28
44	−4	45	72	−69	−119	57	27	−47
45	22	−68	57	−74	−42	16	−10	−61
46	50	−81	47	−67	55	14	−25	−58
47	17	−65	39	17	109	−14	−28	32
48	−53	−26	−16	43	60	5	37	130
49	−52	0	−91	−11	25	−10	82	91
50	6	66	−98	−32	−26	19	49	3

Table 9. *Series* 7

No. of term	0+	50+	100+	150+	200+	250+	300+	350+	400+	450+
1	−0·060	0·666	−0·638	1·196	1·070	−0·267	0·593	−0·726	−2·320	−0·871
2	−0·581	1·744	0·390	1·888	1·588	0·457	−1·914	−0·345	−1·457	−0·826
3	1·517	1·831	0·718	1·370	−1·501	0·847	−2·208	1·438	−1·558	0·388
4	2·200	1·660	−0·181	1·472	0·369	−0·163	−1·316	3·073	−1·738	1·539
5	2·210	1·158	−0·933	3·308	−0·565	−0·627	−3·207	1·964	−0·535	2·453
6	1·513	0·201	−0·870	2·660	−2·778	−1·021	−2·072	2·521	−1·119	−0·032
7	1·519	0·967	−0·156	2·195	−3·380	−1·229	−2·587	1·461	−0·884	−0·085
8	1·156	−0·084	−0·428	1·974	−1·469	−2·485	−2·661	1·322	0·314	0·778
9	2·710	0·600	−1·400	2·882	0·090	−0·924	−2·419	−0·400	0·261	2·117
10	1·421	−1·186	−1·017	4·137	1·250	0·375	−2·857	−0·422	−1·092	1·108
11	0·197	−0·960	−0·796	4·284	2·660	0·381	−2·859	−0·501	−0·420	0·236
12	0·762	0·291	−0·574	3·395	2·309	1·351	−2·526	−1·049	−0·429	0·302
13	0·774	0·818	−0·576	1·606	0·322	2·567	−3·180	−0·095	−1·507	1·369
14	−0·429	2·450	−1·157	0·909	0·735	3·124	−3·628	−0·175	−0·230	0·516
15	−1·001	1·632	−0·970	−0·456	0·804	2·795	−1·725	−1·338	−0·976	1·227
16	−1·505	3·047	−1·761	−0·061	−0·463	2·705	−3·168	−0·590	−1·329	2·700
17	−0·686	2·402	−1·200	−0·182	−0·782	2·730	−3·724	0·443	−1·928	3·006
18	−1·027	1·517	−0·892	−0·567	−0·453	2·860	−4·270	−1·260	−1·959	3·701
19	−1·518	2·550	0·920	−0·180	−0·916	3·655	−4·656	−0·611	0·487	3·070
20	−1·732	1·437	0·704	0·098	0·466	2·883	−5·368	−1·042	0·382	1·421
21	−2·439	1·693	−0·279	1·630	0·572	3·920	−5·492	−0·932	0·221	1·348
22	−1·696	2·455	−0·694	1·039	−0·208	4·688	−5·624	−0·868	−0·737	−0·587
23	−0·791	1·207	−0·879	−0·307	0·707	3·837	−3·301	−1·788	−1·310	−0·459
24	−2·302	0·219	0·840	−1·326	−0·834	4·721	−5·446	0·487	−0·404	−0·667
25	−2·009	0·204	0·564	−1·243	−1·002	2·830	−5·222	0·907	−2·243	−2·541
26	−3·029	−0·222	−0·066	0·119	−1·140	2·901	−2·483	0·802	−0·762	−0·844
27	−2·445	0·350	2·557	−0·874	−2·041	4·371	−2·210	0·198	−1·472	−0·562
28	−1·945	0·059	1·643	−0·769	−2·802	4·196	−2·266	−0·883	−0·560	−0·322
29	−0·863	0·621	1·857	−2·196	−3·613	4·025	−2·122	−1·375	−1·942	−3·020
30	−0·970	0·727	1·845	−1·588	−3·191	3·240	−1·132	−2·324	−2·898	−1·927
31	0·290	0·272	3·143	−2·759	−3·016	3·521	−0·628	−3·390	−2·032	0·157
32	−0·336	1·699	3·554	−2·583	−2·369	4·316	0·789	−2·437	−1·343	1·206
33	1·065	1·850	2·232	−2·503	−2·386	−0·016	0·235	−2·506	−1·413	0·900
34	−0·119	2·031	2·117	−0·830	−2·455	−0·911	−0·012	−2·273	−1·632	3·060
35	1·189	2·443	2·630	0·726	−2·989	−1·524	−0·366	−0·022	−1·451	2·774
36	1·595	3·005	2·868	0·070	−1·781	−1·730	−1·559	0·643	−1·438	1·206
37	1·717	3·491	2·357	0·616	−1·725	−1·512	−2·165	0·323	−1·792	2·438
38	2·729	1·871	2·747	0·315	−1·208	−3·368	−2·913	−0·273	−3·068	1·523
39	2·827	2·177	2·955	−0·533	−1·477	−3·442	−3·367	−1·836	−2·774	1·183
40	2·764	0·790	1·470	−0·148	−2·403	−1·018	−3·359	−1·909	−1·105	0·046
41	2·198	1·113	1·915	−0·782	−1·594	−0·493	−2·111	−0·843	−1·297	−1·493
42	1·975	0·240	0·587	0·063	−1·777	1·487	−3·012	−1·482	0·207	−3·627
43	0·806	0·219	0·549	−0·260	−2·580	0·369	−3·768	−2·331	−0·611	−2·872
44	0·179	−1·257	1·425	−0·789	−3·576	1·709	−4·300	−1·529	−0·098	−1·473
45	0·458	−1·836	1·062	0·014	−3·619	0·459	−3·123	−3·324	1·713	−1·400
46	0·673	−1·174	0·737	−2·347	−3·398	−0·109	−2·664	−3·234	2·422	−0·837
47	0·290	−1·028	2·308	−1·584	−3·558	0·945	−1·828	−3·137	1·694	−1·626
48	1·710	−0·752	2·245	−2·549	−2·666	1·447	0·018	−3·296	1·604	−1·954
49	1·144	−0·989	1·750	−3·155	−1·126	2·865	−1·978	−0·889	0·154	−0·386
50	0·616	−0·261	1·813	−0·893	−0·485	2·284	−1·014	−1·949	−1·390	−0·057

The 501st and 502nd terms are 0·175 and −1·299 respectively.

Table 10. *Series* 8

No. of term	0+	50+	100+	150+	200+	250+	300+	350+	400+	450+
1	0·91	3·81	1·20	3·54	2·20	2·59	−3·98	0·82	−3·62	−2·44
2	3·49	4·38	1·12	3·40	1·57	1·93	−6·05	4·63	−3·84	0·83
3	5·59	4·07	−0·30	5·27	0·06	0·20	−7·88	6·65	−2·95	4·59
4	5·92	2·49	−1·76	6·76	−3·50	−1·77	−7·71	7·52	−2·44	4·60
5	5·24	1·67	−1·94	7·00	−7·26	−3·27	−7·13	6·41	−2·10	2·68
6	3·95	0·51	−1·69	6·29	−7·70	−5·20	−6·65	4·61	−0·77	1·43
7	4·44	0·32	−2·28	6·30	−4·76	−5·01	−6·17	1·47	0·46	2·35
8	4·33	−1·08	−2·68	7·92	−0·13	−2·53	−6·32	−1·11	−0·20	2·98
9	2·74	−2·31	−2·61	9·85	4·90	0·10	−6·72	−2·46	−0·87	2·34
10	1·61	−1·71	−2·10	10·27	7·76	2·73	−6·76	−3·20	−1·29	1·38
11	1·18	0·09	−1·58	7·98	6·41	5·52	−7·26	−2·38	−2·49	1·72
12	0·06	3·41	−1·85	4·55	3·91	7·83	−8·23	−1·20	−2·32	1·72
13	−1·52	5·33	−2·21	0·56	1·90	8·65	−7·15	−1·46	−2·29	2·26
14	−3·21	7·21	−3·27	−1·72	−0·33	8·30	−6·92	−1·60	−2·68	4·32
15	−3·46	7·67	−3·69	−2·35	−2·09	7·54	−7·76	−0·59	−3·74	6·63
16	−3·22	6·35	−3·32	−2·30	−2·59	7·00	−9·35	−1·10	−4·73	8·84
17	−3·34	5·70	−0·88	−1·53	−2·72	7·59	−11·06	−1·53	−2·84	9·47
18	−3·79	4·53	1·39	−0·44	−1·23	7·73	−12·86	−2·18	−0·38	7·42
19	−4·94	3·83	1·69	1·92	0·58	8·63	−14·11	−2·56	1·22	4·78
20	−5·23	4·40	0·47	3·36	1·04	10·31	−14·71	−2·59	0·80	0·96
21	−4·08	4·13	−1·21	2·44	1·57	10·87	−12·43	−3·36	−1·04	−1·80
22	−4·17	2·57	−0·72	−0·33	0·37	11·52	−11·76	−1·91	−1·95	−3·12
23	−4·56	0·96	0·37	−2·82	−1·38	10·07	−11·95	0·48	−3·87	−5·08
24	−5·96	−0·45	0·70	−2·82	−2·84	8·22	−9·74	2·29	−4·04	−4·87
25	−6·72	−0·62	3·15	−2·57	−4·48	9·37	−6·95	2·48	−3·98	−3·38
26	−6·36	−0·40	4·75	−2·18	−6·31	10·40	−5·04	0·70	−2·92	−1·60
27	−4·50	0·49	5·51	−3·31	−8·31	10·78	−4·19	−1·85	−3·16	−3·10
28	−2·74	1·47	5·53	−4·14	−9·18	9·90	−3·22	−4·70	−4·92	−4·53
29	−0·47	1·64	6·47	−5·66	−8·96	9·02	−2·08	−7·64	−5·86	−3·28
30	0·51	2·77	7·91	−6·74	−7·63	9·29	0·12	−8·49	−5·33	−0·14
31	1·87	4·09	7·69	−7·08	−6·30	5·69	1·40	−8·02	−4·35	2·39
32	1·68	5·14	6·63	−5·25	−5·57	0·71	1·47	−6·85	−3·75	5·76
33	2·10	6·06	6·07	−1·51	−5·97	−3·59	0·55	−3·55	−3·40	7·91
34	3·07	7·09	6·23	1·03	−5·56	−6·04	−1·69	0·17	−3·30	7·03
35	4·04	8·27	6·18	2·51	−4·86	−6·35	−4·30	2·28	−3·73	6·22
36	5·64	7·42	6·43	2·56	−3·77	−7·34	−6·80	2·15	−5·52	4·84
37	7·01	6·20	6·93	1·03	−3·20	−8·34	−8·69	−0·61	−6·98	3·40
38	7·66	3·90	5·88	−0·30	−4·03	−6·52	−9·52	−3·65	−6·02	1·37
39	7·11	2·31	4·92	−1·62	−4·43	−3·50	−8·24	−4·56	−4·43	−1·69
40	5·97	0·82	3·06	−1·57	−4·64	0·90	−7·31	−4·67	−1·66	−6·17
41	3·82	−0·03	1·45	−1·18	−5·46	3·11	−7·69	−5·19	−0·22	−8·81
42	1·39	−1·70	1·49	−1·30	−7·27	4·68	−9·11	−4·90	0·49	−8·08
43	0·08	−3·69	1·98	−0·83	−8·88	4·05	−9·29	−6·12	2·36	−5·88
44	0·07	−4·39	2·17	−2·61	−9·53	2·01	−8·33	−7·52	4·77	−3·27
45	0·32	−4·01	3·70	−4·04	−9·60	1·13	−6·35	−8·34	5·77	−2·28
46	2·03	−2·97	5·23	−5·69	−8·46	1·68	−2·80	−8·72	5·56	−2·83
47	3·22	−2·25	5·66	−7·39	−5·63	4·15	−1·88	−6·31	3·39	−2·36
48	3·14	−1·25	5·42	−6·18	−2·45	6·01	−1·69	−4·53	−0·45	−1·24
49	2·51	−0·89	4·33	−2·03	−0·15	5·13	−1·64	−4·15	−3·05	−0·01
50	2·94	0·04	3·94	2·44	1·52	0·72	−1·30	−3·75	−3·96	−0·69

Table 11. *Series* 9

No. of term	0+	50+	100+	150+	200+	250+	300+	350+	400+	450+
1	2·390	−1·093	−0·390	−1·528	1·023	−1·549	−0·093	−0·119	−0·069	0·177
2	0·985	−1·147	−1·119	−1·507	−0·098	−2·845	−1·162	−0·789	0·054	−0·044
3	−0·655	0·472	1·477	−1·378	−0·499	−0·766	0·734	−0·560	−0·903	−0·493
4	−0·679	0·995	3·374	−0·164	−0·493	−0·598	−0·108	−2·073	−0·603	1·064
5	−0·044	0·400	3·173	0·405	1·325	0·223	0·755	−2·348	−0·691	0·688
6	−1·457	−0·838	2·029	1·080	−0·915	−0·475	−0·022	−1·104	−1·737	0·558
7	−0·731	−1·760	1·895	1·358	1·159	−0·521	1·113	−1·271	−0·908	−0·235
8	−0·724	−1·660	−0·669	0·847	0·391	0·798	−1·301	−0·404	0·731	−1·151
9	−1·567	0·058	−1·578	2·333	−0·093	1·567	1·160	0·481	1·316	−1·803
10	−1·654	−0·226	−0·457	2·121	−1·850	0·929	1·397	0·588	−0·432	−3·082
11	−2·416	0·303	0·875	0·822	−1·059	1·927	1·857	0·865	0·241	−1·691
12	−2·821	−0·742	−0·377	1·245	−1·014	0·203	1·502	1·798	0·387	−1·467
13	−0·701	−2·360	0·529	2·618	−0·112	−1·467	0·882	0·663	−0·414	−2·607
14	−1·515	−1·644	2·016	3·134	0·992	−0·550	1·052	0·519	1·696	−2·160
15	−2·112	−1·397	1·186	1·582	0·913	−0·411	2·911	−0·064	1·998	−1·505
16	−1·602	−1·792	−0·430	1·099	2·637	1·384	3·643	0·133	0·263	−0·722
17	−1·805	0·826	−0·832	−0·230	1·982	0·591	1·186	0·441	1·270	−0·876
18	−1·624	1·679	−0·309	1·113	1·315	0·809	1·862	−0·974	0·485	1·045
19	−1·060	0·732	−1·515	0·289	1·054	1·306	0·112	−2·054	−1·974	0·950
20	−0·022	1·093	−1·868	−0·153	0·423	3·434	1·032	0·386	−2·867	0·683
21	0·546	1·058	−2·501	−1·891	−0·181	3·421	−0·491	0·826	−2·328	−0·594
22	−0·886	1·536	−2·059	−2·573	−0·464	2·275	1·032	−1·679	−1·682	−2·596
23	−1·321	0·244	−1·121	−1·459	−0·476	2·355	2·384	−1·119	1·102	−1·126
24	−1·014	0·210	−2·314	0·245	1·183	1·366	0·624	−1·946	2·350	−0·665
25	−2·254	−0·694	−3·101	−0·027	0·973	1·552	−0·917	−1·392	1·344	−0·823
26	0·582	−1·028	−2·921	1·390	0·801	−0·026	−1·363	−0·476	1·133	−0·469
27	0·272	−2·704	−1·972	0·253	−1·514	0·540	−1·593	−0·341	1·642	−2·195
28	0·358	−1·164	−3·015	−1·583	−0·455	0·352	−1·606	−0·470	0·800	0·448
29	0·981	−1·382	−4·109	−1·577	0·204	0·665	−3·104	−0·384	1·977	−0·573
30	−0·497	−1·815	−2·635	−2·094	−0·132	−0·531	−3·599	1·604	0·233	−0·957
31	−1·078	0·006	−1·375	−0·473	0·512	−0·877	−1·665	2·262	−0·732	0·114
32	−0·318	1·277	−2·038	−0·485	0·550	0·195	1·684	1·352	0·864	−1·102
33	−0·597	1·496	−0·713	−2·421	0·717	0·565	2·532	0·700	0·571	−1·724
34	1·697	1·785	−0·523	0·261	0·132	0·470	0·471	1·276	1·853	−1·961
35	2·585	1·035	−0·443	1·019	0·400	0·352	−1·502	1·128	1·589	−1·539
36	0·170	1·615	−1·932	1·259	−0·241	1·325	−1·763	1·682	0·980	−1·111
37	0·497	1·734	−1·592	2·917	0·136	1·383	−2·083	−1·061	1·530	−2·282
38	0·437	3·754	0·252	1·018	1·115	0·956	−2·674	0·442	1·110	−0·783
39	1·554	1·252	0·527	1·601	−0·044	1·191	−2·365	0·482	−0·024	−0·421
40	1·474	1·540	−0·438	0·989	0·451	0·895	−2·846	0·395	0·693	−2·011
41	−0·809	0·808	0·644	−0·188	0·503	0·978	−4·160	0·722	0·102	−1·903
42	0·240	1·408	0·095	−0·910	0·215	0·707	−3·199	0·477	1·465	−2·617
43	1·088	−0·382	0·452	0·166	−0·392	−0·319	−2·662	0·646	0·086	−2·433
44	−0·609	−0·069	0·955	−0·445	2·271	0·747	−4·524	1·893	0·255	−2·710
45	0·440	−0·179	1·307	−0·126	3·453	−1·327	−3·303	2·771	0·186	−2·601
46	0·486	−1·966	1·319	−1·096	2·055	1·291	−1·198	1·159	1·571	−2·670
47	−0·523	1·004	−0·085	−0·008	2·397	−0·848	−0·765	−1·065	−0·233	−0·866
48	−1·337	0·701	−0·271	−0·021	2·095	−0·142	0·347	−2·119	−0·151	−0·016
49	−1·116	1·259	−1·908	2·390	1·924	−0·749	0·432	−1·117	−0·485	0·940
50	−1·675	1·385	−1·725	2·176	1·331	0·167	0·961	−2·044	0·596	0·001

The 501st and 502nd terms are 0·950 and 3·505 respectively.

Table 12. *Series* 10

No. of term	0+	50+	100+	150+	200+	250+	300+	350+	400+	450+
1	−0·767	−1·484	0·004	−4·339	0·422	−6·040	−0·388	−1·455	−0·607	−0·268
2	−2·015	1·188	3·435	−2·353	−1·673	−4·886	0·067	−3·827	−0·394	0·644
3	−1·877	2·449	6·949	−0·014	−0·726	−2·132	1·022	−5·830	−0·820	1·531
4	−2·514	1·262	7·956	2·241	−0·877	−0·377	1·069	−5·604	−2·443	1·920
5	−2·558	−1·597	7·172	3·830	0·557	0·130	1·778	−4·520	−3·185	1·111
6	−2·281	−4·047	3·242	3·940	1·443	1·130	0·120	−2·574	−1·551	−0·888
7	−2·797	−3·596	−1·598	4·751	1·215	2·745	0·403	−0·091	1·202	−3·336
8	−3·590	−2·158	−3·835	5·378	−1·235	3·383	1·781	1·776	1·666	−6·307
9	−4·967	−0·273	−2·545	4·362	−3·025	4·276	3·614	2·863	1·473	−6·961
10	−6·489	0·037	−1·259	3·354	−3·724	3·215	4·587	4·060	1·174	−5·971
11	−5·356	−2·183	0·417	4·127	−2·696	−0·068	4·121	3·697	0·141	−5·694
12	−4·162	−4·064	3·104	5·996	−0·112	−2·233	3·291	2·556	1·264	−5·438
13	−4·012	−4·776	4·392	6·115	2·138	−2·833	4·471	0·899	3·318	−4·640
14	−3·934	−5·013	2·849	4·827	5·045	−0·616	6·916	−0·156	3·281	−3·107
15	−4·127	−2·301	0·106	2·022	6·462	1·330	6·558	−0·180	3·220	−1·974
16	−4·196	1·655	−1·617	0·924	5·901	2·580	5·618	−1·094	2·386	0·428
17	−3·613	3·703	−3·347	0·294	4·314	3·479	3·013	−3·168	−0·959	2·407
18	−1·898	4·339	−4·741	−0·291	2·218	5·971	1·537	−2·551	−5·115	3·117
19	0·265	3·979	−6·043	−2·359	0·102	8·250	−0·307	−0·400	−7·475	1·631
20	0·354	3·744	−6·336	−5·022	−1·461	8·364	−0·074	−0·840	−7·347	−2·360
21	−1·064	2·373	−5·069	−5·804	−2·134	7·431	2·456	−1·844	−3·242	−4·538
22	−2·361	0·948	−4·722	−3·628	−0·434	5·358	3·363	−3·555	2·457	−4·477
23	−4·320	−0·838	−5·761	−1·116	1·563	3·730	1·554	−4·380	5·668	−3·478
24	−2·989	−2·423	−6·897	1·976	2·737	1·398	−1·335	−3·517	6·139	−2·057
25	−0·856	−4·951	−6·678	2·985	0·715	0·213	−3·839	−2·019	5·561	−2·718
26	0·911	−5·398	−6·913	0·712	−1·037	−0·113	−5·161	−0·933	3·848	−1·514
27	2·411	−4·845	−8·374	−2·286	−1·294	0·434	−6·862	−0·400	3·429	−0·879
28	1·700	−4·445	−8·390	−4·965	−1·037	0·003	−8·566	1·630	2·081	−1·167
29	−0·414	−2·461	−6·417	−4·791	0·018	−1·091	−7·657	4·255	−0·158	−0·730
30	−1·623	0·792	−4·902	−3·273	1·089	−1·006	−2·456	5·218	−0·350	−1·322
31	−2·176	3·598	−2·896	−3·626	1·905	0·003	3·659	4·312	0·265	−2·813
32	0·116	5·347	−1·258	−2·091	1·684	0·977	5·724	3·410	2·320	−4·394
33	3·800	5·117	−0·379	0·532	1·299	1·425	2·965	2·723	4·008	−4·966
34	4·292	4·571	−1·720	2·890	0·346	2·404	−1·364	2·973	4·229	−4·377
35	3·318	4·203	−3·294	5·830	−0·133	3·315	−5·066	0·847	4·178	−4·613
36	1·941	6·092	−2·512	5·986	0·796	3·400	−7·564	−0·112	3·591	−3·669
37	2·030	5·852	−0·589	5·271	0·898	3·274	−8·153	−0·065	1·837	−2·151
38	2·737	4·931	0·170	3·794	1·041	2·796	−8·032	0·380	0·919	−2·542
39	1·186	3·306	1·126	1·350	1·199	2·417	−8·919	1·172	0·194	−3·624
40	0·177	2·579	1·248	−1·322	1·013	1·967	−8·994	1·577	1·219	−5·332
41	0·689	0·802	1·262	−1·963	0·123	0·637	−8·096	1·794	1·330	−6·487
42	0·061	−0·476	1·719	−1·944	1·900	0·464	−8·932	3·078	1·108	−7·179
43	0·162	−1·104	2·567	−1·282	5·481	−1·135	−9·081	5·260	0·740	−7·255
44	0·634	−2·942	3·283	−1·535	7·134	−0·190	−6·721	5·406	1·831	−7·061
45	0·093	−1·681	2·243	−1·055	7·504	−0·489	−3·617	2·252	1·411	−5·005
46	−1·551	0·324	0·555	−0·414	6·782	−0·585	−0·272	−2·345	0·486	−1·992
47	−2·869	2·455	−2·419	2·462	5·633	−1·148	1·942	−4·823	−0·656	1·252
48	−4·055	3·924	−4·664	5·091	4·136	−0·803	3·233	−6·176	−0·369	2·374
49	−4·119	2·699	−5·448	5·392	0·184	−0·403	2·466	−4·452	0·100	2·935
50	−3·651	−0·113	−5·168	3·288	−4·711	−1·203	0·307	−1·755	0·250	5·547

Table 13. *Series* 11

No. of term	0+	50+	100+	150+	200+	250+	300+	350+	400+	450+
1	−1·301	1·170	0·612	−1·339	−2·257	2·498	1·177	−0·871	0·363	−0·076
2	−1·335	0·075	−0·439	−1·221	−1·741	2·166	−0·081	−1·167	0·620	0·114
3	−1·576	−0·486	−0·365	−0·445	−0·616	0·982	−0·848	−1·905	0·562	−0·389
4	−1·122	−1·443	−0·782	−0·673	0·852	0·843	−2·310	−2·071	0·961	−0·221
5	0·162	1·165	−0·081	−0·640	−1·186	1·785	−0·003	0·002	−0·628	−1·452
6	−0·452	0·808	−1·575	0·642	−1·173	1·081	0·674	−1·759	0·157	0·848
7	−1·075	−0·228	0·026	−0·339	−0·415	0·092	0·801	−0·252	−0·138	1·904
8	−1·030	−0·816	−1·066	1·731	2·073	0·293	0·840	0·112	0·438	0·860
9	0·178	−2·128	0·996	−0·323	0·656	2·189	0·029	1·907	1·097	0·145
10	0·231	−2·151	−0·597	1·207	1·076	1·534	−0·001	0·320	0·839	0·763
11	−0·580	−0·541	−0·675	0·795	−0·571	1·171	−0·086	−1·163	−0·418	0·706
12	−2·478	1·039	1·200	−1·024	2·018	0·508	−1·190	1·378	0·978	−0·547
13	1·091	−1·203	0·101	−1·198	0·880	0·346	−0·452	2·367	−0·615	−0·550
14	2·273	−0·468	0·470	−1·520	−0·293	−1·557	−0·444	0·523	−1·122	−0·300
15	2·614	−0·012	1·183	−1·053	−0·092	0·285	−0·748	−0·309	−2·996	0·079
16	2·545	−1·234	−0·604	1·019	−0·271	1·342	−0·100	−0·759	−2·796	−0·427
17	2·644	0·307	−0·125	0·336	0·069	1·662	−0·122	−0·571	−0·796	−0·293
18	0·085	−0·513	−0·464	1·967	−0·784	2·921	0·255	−0·748	−0·130	1·119
19	−1·679	−0·283	−0·774	−1·686	−0·532	2·086	−0·215	0·012	−1·669	−0·221
20	−1·041	0·141	−1·469	−0·291	−1·862	2·610	−0·355	−0·715	−0·956	0·345
21	−1·369	0·318	−1·032	1·022	−2·056	2·528	1·874	−0·122	−1·522	−0·904
22	−1·927	1·169	−0·478	1·817	−0·931	2·307	−0·111	−0·551	−0·367	−0·002
23	−1·462	0·416	−2·042	2·689	−1·033	1·079	−1·866	−0·817	0·498	−0·482
24	−1·523	−0·315	−0·420	1·226	1·356	−0·187	−0·707	−0·005	2·575	0·672
25	−0·286	0·372	1·586	1·083	0·838	0·323	−0·426	0·737	0·999	−0·076
26	0·343	−0·913	−0·867	−0·299	1·005	1·073	1·126	−0·434	2·012	0·141
27	−2·558	1·618	0·467	−0·412	0·847	0·085	0·587	1·337	−0·617	−0·264
28	0·731	0·742	−0·360	1·914	−0·706	0·225	1·278	−0·089	0·130	1·940
29	−0·630	0·217	2·127	1·122	−0·410	0·926	0·386	−0·837	1·985	0·503
30	0·857	0·979	0·979	−1·742	−0·393	1·235	−0·516	0·720	0·764	−0·580
31	0·656	−0·388	−0·222	−1·145	−0·593	0·453	−1·360	−0·809	−0·500	−2·443
32	0·062	−1·779	−1·313	0·388	0·206	0·904	1·101	1·226	0·161	0·028
33	−0·304	0·743	−0·475	2·669	0·808	2·384	1·930	0·356	−0·785	1·370
34	−0·244	−0·113	−0·786	1·474	1·058	0·791	0·486	0·785	0·511	2·055
35	1·079	−1·830	0·682	−0·345	−0·703	1·303	0·209	−1·835	0·778	0·300
36	1·140	−2·154	0·730	−1·160	0·161	0·368	1·173	−0·148	−0·235	0·360
37	−0·457	−1·251	1·439	−0·987	0·434	−0·455	1·895	0·250	−0·561	−1·062
38	−0·554	0·220	0·727	−0·968	0·638	−0·425	1·273	−0·935	−0·141	−0·498
39	0·698	0·369	0·191	−2·854	0·614	−2·542	0·729	0·602	−0·765	0·593
40	1·176	−1·228	−1·452	−0·357	−0·276	−1·284	−0·593	0·796	−1·177	−0·211
41	0·923	−0·925	−1·934	−0·483	−0·619	−2·451	0·803	0·428	−0·863	−0·009
42	−1·984	0·704	−0·920	−1·678	−1·858	−1·428	0·333	−0·296	0·010	−0·703
43	−1·603	0·205	0·858	−1·335	−0·811	0·169	0·523	0·092	−0·620	−0·171
44	−2·393	−1·136	0·385	0·588	−0·414	1·360	0·492	0·615	0·385	0·816
45	−2·101	−0·915	−0·658	0·298	−0·257	2·134	0·701	−0·293	−0·625	0·907
46	−1·377	−0·260	0·485	−0·525	1·004	−0·035	−0·184	−1·405	1·051	0·758
47	−0·525	0·373	−0·035	−0·599	2·024	−0·553	−1·744	−2·676	1·411	1·138
48	−0·986	0·776	1·162	−0·978	2·732	1·849	−0·358	−0·222	1·312	0·329
49	−0·619	0·925	−0·216	−0·110	2·918	1·829	−1·500	−0·286	0·282	−0·563
50	−1·290	−0·811	−2·405	−1·070	3·102	1·091	−2·366	−0·778	−0·749	1·225

The 501st and 502nd terms are 0·501 and 0·445 respectively.

Table 14. *Series* 12

No. of term	0+	50+	100+	150+	200+	250+	300+	350+	400+	450+
1	−2·394	0·486	−0·917	−3·592	−4·036	4·305	−0·366	−3·727	2·109	−1·211
2	−3·088	−1·386	−1·734	−2·264	−1·176	2·207	−3·694	−4·299	2·603	−1·004
3	−2·038	−0·603	−1·530	−1·334	−0·462	2·060	−3·884	−2·863	1·181	−1·951
4	−1·150	0·838	−2·391	0·306	−1·093	2·243	−1·751	−2·759	0·155	−0·796
5	−1·321	0·995	−1·839	0·665	−1·386	1·530	0·817	−1·855	−0·558	2·004
6	−1·908	−0·140	−1·893	2·309	1·095	0·854	2·614	−0·549	−0·254	3·462
7	−1·260	−2·780	−0·167	1·885	2·553	2·364	2·496	2·230	1·097	3·952
8	−0·202	−5·139	0·166	2·126	3·337	3·707	1·438	3·048	2·173	3·379
9	−0·172	−4·804	−0·409	2·191	1·823	4·067	0·247	1·075	1·424	2·447
10	−2·566	−1·676	0·667	0·323	2·355	3·128	−1·637	1·036	1·457	0·455
11	−1·646	−0·644	1·039	−1·938	2·559	1·753	−2·376	2·970	0·276	−1·273
12	1·746	−0·339	1·280	−3·813	1·344	−1·192	−2·239	3·271	−1·547	−1·928
13	5·357	−0·063	2·071	−4·279	0·107	−1·903	−2·023	1·805	−4·836	−1·405
14	7·565	−1·134	1·034	−1·781	−0·825	−0·155	−1·206	−0·409	−7·342	−1·009
15	8·287	−0·909	−0·023	0·516	−0·892	2·443	−0·437	−1·924	−6·454	−0·700
16	5·418	−0·946	−1·006	3·426	−1·353	5·686	0·377	−2·659	−3·559	0·853
17	0·138	−0·869	−1·869	1·824	−1·574	7·119	0·419	−1·952	−2·357	1·068
18	−3·599	−0·342	−3·022	0·003	−2·917	7·598	−0·083	−1·532	−1·769	1·093
19	−5·397	0·376	−3·422	0·113	−4·478	7·326	1·573	−0·832	−2·289	−0·236
20	−6·064	1·754	−2·731	1·940	−4·398	6·567	1·661	−0·700	−2·001	−0·808
21	−5·434	2·157	−3·335	4·766	−3·632	4·640	−0·825	−1·171	−0·558	−1·253
22	−4·468	1·181	−2·723	5·499	−0·440	1·633	−2·446	−0·943	2·961	−0·302
23	−2·484	0·592	0·258	4·749	2·170	−0·201	−2·703	0·285	4·536	0·218
24	−0·156	−0·852	0·779	2·175	3·612	0·036	−0·625	0·351	5·521	0·532
25	−1·487	0·385	1·194	−0·394	3·735	0·225	1·251	1·581	3·188	0·212
26	−0·827	1·591	0·565	0·393	1·597	0·454	2·967	1·474	0·876	1·907
27	−0·796	1·775	2·151	1·752	−0·521	1·313	3·024	−0·006	1·355	2·495
28	0·395	2·136	3·063	−0·012	−1·765	2·453	1·327	−0·023	1·816	1·211
29	1·488	1·074	2·072	−2·034	−2·274	2·494	−1·412	−0·832	0·821	−2·359
30	1·502	−1·666	−0·566	−1·843	−1·413	2·421	−1·116	0·323	0·155	−3·172
31	0·604	−1·626	−2·133	1·658	0·391	3·800	1·409	1·127	−1·024	−0·940
32	−0·331	−1·069	−2·850	4·220	2·194	3·761	2·593	1·863	−0·693	2·697
33	0·413	−2·193	−1·386	3·468	1·515	3·540	2·358	−0·349	0·527	3·638
34	1·760	−4·032	0·630	0·545	0·731	2·381	2·470	−1·463	0·692	3·058
35	1·272	−4·589	2·825	−2·122	0·480	0·395	3·433	−1·185	−0·064	0·483
36	−0·034	−2·813	3·520	−3·574	0·801	−1·182	3·814	−1·507	−0·557	−1·496
37	0·024	−0·430	2·659	−5·725	1·255	−4·039	3·208	−0·463	−1·346	−1·294
38	1·220	−0·295	−0·297	−4·867	0·704	−5·136	1·029	1·040	−2·379	−0·886
39	2·253	−1·034	−3·586	−2·975	−0·472	−6·081	0·331	1·804	−2·807	−0·337
40	−0·116	−0·286	−4·716	−2·516	−2·729	−5·549	0·182	1·168	−1·888	−0·631
41	−2·857	0·407	−2·537	−2·616	−3·577	−2·895	0·558	0·475	−1·294	−0·696
42	−5·478	−0·545	−0·047	−1·031	−2·984	0·951	1·015	0·554	−0·094	0·366
43	−6·698	−1·718	0·558	0·472	−1·751	4·627	1·538	0·078	−0·081	1·657
44	−6·006	−1·877	1·123	0·509	0·570	4·579	1·001	−1·596	1·008	2·398
45	−3·783	−0·833	0·921	−0·275	3·526	2·171	−1·412	−4·470	2·561	2·947
46	−2·144	0·798	1·614	−1·535	6·326	1·947	−2·412	−4·342	3·625	2·372
47	−1·086	2·220	1·099	−1·661	8·114	2·886	−3·447	−2·827	2·989	0·573
48	−1·413	1·232	−2·004	−2·130	8·864	3·292	−4·952	−1·717	0·726	0·669
49	0·159	0·857	−4·092	−3·769	8·191	3·355	−4·594	−0·112	−0·772	0·950
50	0·956	−0·112	−4·721	−4·822	6·745	1·964	−3·745	1·355	−1·098	1·156

Table 15. *Series* 13

No. of term	0+	50+	100+	150+	200+	250+	300+	350+	400+	450+
1	0·032	−1·048	−0·765	0·828	0·383	−0·654	0·359	−0·690	1·494	−0·145
2	−2·155	−1·443	0·566	0·326	1·805	0·163	−0·179	0·798	−1·191	−1·625
3	−3·152	0·399	0·072	−1·922	1·010	0·694	−0·132	0·722	−0·488	0·857
4	0·009	0·484	−0·942	−2·452	1·934	0·945	−0·240	2·344	−0·312	−0·305
5	−0·426	−0·664	−0·018	−0·442	1·594	−1·615	−0·630	1·154	0·981	−0·026
6	0·773	−0·662	0·176	−0·092	1·793	−0·217	−0·232	−0·847	0·243	0·539
7	−1·167	−0·506	−0·565	−0·989	1·959	−1·432	−0·834	−0·248	0·343	−0·415
8	0·108	−2·637	0·031	−0·611	2·169	−0·575	−1·388	−1·205	−0·117	0·715
9	1·013	−0·423	0·868	0·106	0·930	0·078	−2·041	−0·805	0·244	1·585
10	0·886	−0·052	0·559	1·231	−0·621	−1·060	0·075	0·476	0·954	1·275
11	−0·553	0·052	0·384	−0·925	0·610	−0·512	−0·228	−1·595	0·899	−0·924
12	−0·302	−0·161	0·984	0·176	0·111	1·145	−1·453	−0·830	0·854	−0·346
13	−0·760	−0·793	−0·465	−0·635	−1·139	−0·693	0·018	−0·791	0·076	−1·822
14	−1·115	1·006	1·961	1·113	1·894	−0·199	0·054	0·125	−1·085	−0·422
15	−0·294	1·312	0·944	−1·338	1·105	0·882	−0·716	−0·816	1·577	−0·204
16	−0·866	1·713	0·547	−2·081	−1·012	−0·047	1·389	−0·201	0·142	0·076
17	−0·393	−0·762	1·554	−1·092	0·290	0·088	−0·748	−0·657	0·081	−0·097
18	0·038	−0·794	0·856	0·678	0·181	1·385	0·240	−0·973	−0·828	0·653
19	0·501	1·331	0·827	−0·471	0·455	−0·073	−0·334	−1·562	0·137	−1·724
20	−0·834	−0·877	0·771	0·465	1·421	−0·526	0·052	−0·360	−0·419	−0·818
21	0·944	−0·785	−0·385	−1·505	2·358	1·473	1·767	−1·676	1·563	−0·372
22	0·350	−1·190	−0·714	−1·078	0·702	−0·168	0·056	−1·126	1·473	−1·206
23	0·153	0·537	−2·336	0·103	2·639	−0·960	−0·934	0·916	−0·426	−0·521
24	−0·385	−0·330	0·564	0·916	1·869	−1·885	−0·555	−0·788	−1·020	−0·440
25	−0·494	0·183	0·544	0·469	0·746	−2·311	1·229	−0·681	1·143	1·015
26	0·228	−2·011	−0·267	2·254	0·311	−2·151	2·391	0·170	−2·405	−0·860
27	0·557	−1·006	−0·476	1·179	0·236	−1·838	−0·853	−0·081	0·090	−0·482
28	−1·857	−2·128	−0·868	0·988	−0·157	−0·173	−0·739	0·179	1·221	0·593
29	0·771	−0·204	−2·337	−0·846	−1·691	−0·668	0·801	1·588	0·378	1·226
30	0·173	−0·158	−0·135	0·701	0·049	0·830	0·690	−0·475	−0·564	−0·245
31	1·272	0·246	0·005	1·607	0·448	0·415	2·299	−0·809	−0·576	0·871
32	0·539	0·285	−1·266	−0·275	−1·719	−0·306	1·615	−0·253	−1·412	2·606
33	1·108	1·060	0·863	−0·396	−2·763	−0·389	0·053	1·141	0·154	0·043
34	−0·860	0·720	0·249	−0·032	−0·371	0·391	−0·730	−0·202	0·624	0·071
35	−0·733	0·112	1·312	0·321	0·277	−0·967	0·228	1·439	1·448	0·896
36	1·139	−0·052	−0·836	1·523	−0·547	−0·429	−0·419	0·451	1·256	−0·023
37	0·630	1·460	0·497	−0·522	−0·638	1·235	0·028	2·862	0·941	−1·338
38	−0·775	0·421	1·012	0·692	−1·001	−0·450	0·096	−0·447	0·345	0·198
39	−0·736	−1·425	0·062	0·989	−0·094	0·269	−0·678	−1·514	0·974	1·789
40	−0·560	−1·781	−0·463	−0·512	0·960	0·168	−0·556	−0·042	1·518	−0·870
41	0·666	−1·169	0·823	−0·427	−0·980	1·129	−1·797	−1·279	0·345	−1·309
42	0·088	−0·867	−0·398	0·033	−0·238	0·144	−1·571	−2·557	−0·789	−0·328
43	0·084	−0·408	−1·662	0·214	−0·783	−0·294	0·393	−1·226	−1·358	0·518
44	0·914	−0·334	0·634	−0·945	−1·547	1·745	0·030	0·176	−0·329	−0·075
45	1·293	−0·255	−0·476	−0·438	−1·631	0·160	0·290	−0·143	0·504	1·289
46	−1·812	−0·488	1·440	−0·973	0·269	−0·536	0·035	−1·317	1·022	−0·383
47	−1·849	−0·454	−0·573	−0·018	1·665	0·470	−2·720	−2·487	0·581	−1·475
48	0·193	0·298	0·080	0·375	−1·277	−1·915	−0·844	−1·157	0·370	−0·770
49	0·590	−0·254	2·936	0·613	−0·830	0·327	−1·109	−0·382	−0·380	−0·573
50	0·172	−0·583	−0·554	1·805	−0·815	1·005	−0·737	0·808	0·541	−1·011

The 501st and 502nd terms are −0·662 and −0·337 respectively.

Table 16. *Series* 14

No. of term	0+	50+	100+	150+	200+	250+	300+	350+	400+	450+
1	−5·539	−1·403	0·053	−1·511	3·610	0·158	0·581	1·721	0·234	−1·300
2	−5·006	−0·205	−0·547	−4·781	3·981	1·858	−0·194	4·336	−0·887	−0·744
3	−3·163	−0·189	−0·646	−4·945	4·169	0·350	−1·134	5·063	−0·112	−0·195
4	−0·204	−0·767	−0·261	−3·141	4·388	−0·761	−1·382	2·554	0·564	0·697
5	0·191	−1·255	−0·529	−1·972	4·701	−2·444	−1·788	0·030	1·019	0·449
6	0·420	−3·634	−0·421	−1·209	5·147	−2·883	−2·663	−2·449	0·722	0·861
7	1·379	−3·793	0·670	−0·238	4·241	−1·871	−4·077	−3·514	0·529	2·307
8	2·193	−2·407	1·506	1·574	1·470	−1·677	−3·078	−2·165	1·175	3·383
9	1·170	−0·699	1·706	0·925	0·107	−1·421	−1·575	−2·219	1·927	1·643
10	−0·112	0·273	2·107	0·407	−0·506	0·421	−1·647	−2·189	2·386	−0·230
11	−1·468	−0·143	1·000	−0·650	−1·750	0·480	−1·006	−2·089	1·737	−2·896
12	−2·674	0·713	2·008	0·195	0·223	0·119	−0·229	−1·079	−0·367	−3·493
13	−2·501	2·167	2·652	−0·799	2·225	0·773	−0·465	−0·958	0·305	−2·598
14	−2·281	3·741	2·461	−3·057	1·324	0·744	0·992	−0·715	0·661	−1·036
15	−1·651	2·269	2·935	−4·055	0·634	0·520	0·576	−0·965	0·656	0·063
16	−0·638	−0·168	2·854	−2·254	0·216	1·585	0·377	−1·677	−0·437	1·240
17	0·625	0·011	2·499	−0·923	0·376	1·411	−0·207	−2·924	−0·672	−0·391
18	0·172	−0·780	2·093	0·577	1·726	0·233	−0·364	−2·738	−0·939	−1·869
19	0·821	−1·649	0·668	−0·409	4·069	1·024	1·470	−3·226	0·866	−2·232
20	1·167	−2·614	−1·026	−1·816	4·315	0·842	1·855	−3·305	2·895	−2·727
21	1·026	−1·514	−3·798	−1·690	5·351	−0·546	0·372	−1·107	2·326	−2·404
22	0·160	−0·688	−3·101	−0·035	5·597	−2·906	−1·074	−0·353	0·091	−1·722
23	−0·831	0·183	−0·968	1·275	4·228	−5·235	−0·138	−0·516	0·080	0·324
24	−0·766	−1·466	0·219	3·674	2·163	−6·456	2·776	−0·221	−2·362	0·357
25	0·130	−2·710	0·249	4·583	0·501	−6·323	2·270	−0·066	−2·549	−0·251
26	−1·331	−4·376	−0·704	4·192	−0·687	−3·900	0·370	0·217	−0·401	0·138
27	−0·758	−3·663	−3·236	1·474	−2·697	−1·796	0·073	1·859	1·211	1·504
28	0·005	−1·999	−3·342	0·226	−2·575	0·804	0·585	1·462	0·969	1·340
29	1·656	−0·122	−2·054	1·119	−1·035	2·197	2·906	−0·731	−0·116	1·593
30	2·359	1·151	−1·854	0·843	−1·571	1·709	4·519	−1·128	−2·024	3·689
31	2·874	2·387	−0·150	−0·029	−3·973	0·392	3·571	−0·034	−2·014	3·304
32	1·122	2·770	1·011	−0·485	−3·956	−0·032	0·939	0·324	−0·580	1·861
33	−0·936	1·966	2·499	−0·198	−2·088	−1·198	−0·525	1·813	1·817	1·291
34	−0·451	0·725	1·408	1·548	−0·866	−1·731	−1·466	2·283	3·545	0·467
35	0·601	1·275	0·796	1·279	−0·547	−0·070	−1·322	4·467	3·932	−1·470
36	0·112	1·461	1·183	1·326	−1·169	0·339	−0·625	3·325	2·898	−1·653
37	−0·913	−0·456	0·966	1·807	−1·107	0·676	−0·705	−0·090	2·195	0·706
38	−1·621	−3·013	0·008	0·813	0·327	0·743	−1·019	−1·803	2·484	0·733
39	−0·660	−4·255	0·349	−0·436	−0·067	1·608	−2·565	−3·218	1·980	−0·856
40	0·172	−4·041	−0·018	−0·853	−0·475	1·541	−3·883	−5·195	0·147	−1·636
41	0·604	−2·726	−1·856	−0·507	−1·272	0·598	−2·596	−5·332	−2·187	−0·854
42	1·492	−1·312	−1·399	−1·076	−2·709	1·632	−0·944	−3·091	−2·808	−0·196
43	2·632	−0·335	−1·087	−1·368	−3·975	1·656	0·550	−0·878	−1·491	1·500
44	0·338	−0·201	0·944	−1·940	−2·749	0·470	1·112	−0·737	0·786	1·365
45	−2·794	−0·507	1·009	−1·468	0·629	0·159	−1·772	−2·859	2·191	−0·723
46	−3·049	−0·160	0·718	−0·270	0·789	−1·975	−3·349	−3·933	2·387	−2·248
47	−1·367	−0·176	3·221	1·050	−0·277	−1·925	−3·907	−3·279	1·150	−2·684
48	0·193	−0·697	2·630	3·095	−1·514	−0·125	−3·360	−0·833	0·613	−2·840
49	−0·152	−1·443	2·111	3·263	−2·181	1·184	−2·433	2·218	−0·046	−2·444
50	−1·707	−0·674	1·333	3·846	−1·479	1·186	−0·198	1·665	−1·982	−1·605

Table 17. *Series* 15

No. of term	0+	50+	100+	150+	200+	250+	300+	350+	400+	450+
1	−0·855	2·075	−1·762	−0·544	−0·486	−0·089	0·764	−0·399	−1·032	0·232
2	1·307	1·113	0·808	0·026	0·014	−0·015	0·770	0·928	−0·561	−1·831
3	1·199	0·287	−0·659	−0·174	−0·677	0·251	−0·794	1·222	−3·251	0·235
4	1·244	0·661	−1·588	−0·002	0·416	−1·055	1·107	−1·566	−0·983	0·538
5	1·259	−0·604	−0·965	0·011	0·586	2·122	0·462	−2·577	0·736	−0·550
6	0·129	−0·867	−1·428	−0·290	0·676	0·638	−0·679	−1·579	0·836	0·762
7	−0·476	−1·817	−0·625	1·032	−0·591	−0·403	1·336	0·023	0·803	−0·953
8	−0·810	0·970	−1·698	1·143	0·472	−0·278	0·236	−0·173	1·048	1·684
9	−0·306	0·773	2·032	1·236	0·668	−1·606	1·207	−0·118	0·987	1·821
10	0·173	0·541	−0·871	0·454	0·049	−1·147	0·007	−0·817	−0·926	−0·083
11	−0·801	0·494	0·096	0·103	0·586	0·005	−1·216	0·997	−0·270	−1·031
12	−0·220	−0·342	−1·203	−0·047	0·165	0·281	−1·249	0·392	0·427	−0·963
13	−1·618	−0·050	0·909	−0·632	1·116	2·150	−1·696	−0·603	−0·255	2·015
14	−1·287	−0·093	0·176	−0·976	0·188	−1·157	0·729	−0·229	−1·159	−0·905
15	−0·032	1·034	1·433	0·606	−0·419	−0·698	−1·579	−1·212	2·104	−1·069
16	−0·571	0·246	−1·273	1·714	0·794	−0·511	1·242	0·506	0·302	0·114
17	1·816	−0·193	−1·877	−0·508	0·574	−0·512	1·704	−1·335	1·022	−0·522
18	0·342	0·507	0·229	1·053	0·977	−0·113	0·160	−0·352	0·073	−0·062
19	0·620	0·325	−0·033	1·193	0·640	−0·801	−1·077	0·851	0·015	−1·121
20	0·406	−0·040	0·274	2·321	0·241	−0·199	−1·159	0·772	0·202	1·217
21	−1·088	0·311	0·724	0·097	−0·602	−0·937	−1·001	−0·080	0·402	1·649
22	−0·166	−0·314	−0·158	0·455	0·525	−1·104	0·559	0·863	0·183	0·027
23	−0·205	0·741	−0·271	−1·979	1·219	1·027	0·154	0·685	0·603	−0·960
24	−0·417	2·126	−0·125	0·621	0·059	−0·287	0·113	−0·363	−0·777	−0·235
25	−0·544	−0·835	1·604	1·357	−0·012	−2·073	1·136	1·064	0·967	−0·435
26	0·651	−1·895	−0·117	0·396	−0·520	0·181	0·254	−1·096	1·182	−0·120
27	0·719	0·037	−0·485	−0·368	1·384	0·460	0·227	−0·316	−0·633	1·201
28	−1·160	−0·068	−0·549	−1·309	1·984	0·371	0·712	−1·149	−1·587	0·145
29	0·396	1·332	−0·237	−0·290	−1·423	−0·399	−0·680	0·313	−1·123	−0·249
30	0·394	0·157	−0·554	−0·102	−1·370	0·395	−0·166	−0·734	0·449	0·608
31	0·460	−0·181	0·569	0·404	−0·077	−0·939	−0·787	1·414	0·564	−0·791
32	0·341	−0·271	1·470	1·916	1·032	0·127	−0·385	−1·084	0·676	2·151
33	−0·549	−0·736	1·052	−0·687	1·015	0·736	0·727	−1·149	−0·275	−0·644
34	−0·536	−1·176	0·127	−2·022	1·232	−0·244	1·055	−0·930	1·308	1·054
35	−1·602	1·663	1·389	−0·914	−0·808	−0·778	−0·370	0·557	−0·578	1·848
36	−0·042	1·545	0·098	−0·759	0·644	−0·283	1·277	0·179	−0·934	1·185
37	−0·013	−0·325	0·356	−0·565	1·860	0·128	0·190	1·018	−1·075	−1·822
38	−0·051	−0·067	−1·255	−1·378	1·998	0·069	−0·472	−0·861	−0·753	−0·462
39	1·128	0·595	0·303	0·052	−0·169	−2·273	0·460	−2·163	0·319	−0·048
40	1·635	0·505	1·444	1·183	1·312	−0·731	0·783	0·756	−0·482	0·163
41	−1·557	−0·639	2·046	1·848	0·656	−1·906	−0·705	−0·647	−1·108	−0·328
42	1·396	0·413	1·345	0·360	1·394	−0·713	1·226	0·732	−1·633	0·104
43	1·783	−0·640	1·117	0·406	0·247	−0·605	1·922	1·447	−0·694	−1·090
44	1·125	−1·148	−1·120	−1·401	0·704	0·142	0·211	1·270	−0·041	0·869
45	1·029	0·518	−0·004	−0·998	−0·223	0·425	0·245	−0·055	−0·029	0·518
46	0·002	−0·148	0·528	−0·699	1·367	−2·728	−1·504	0·964	−0·620	−0·914
47	0·352	−0·109	1·131	−0·813	−0·926	−1·953	0·369	−0·373	−0·169	0·054
48	1·398	−1·018	0·945	−0·286	−2·743	−0·218	−0·853	−1·207	−0·007	−1·359
49	0·329	−0·656	0·850	−0·084	0·452	0·849	0·308	0·902	0·456	−0·005
50	−0·502	0·018	0·285	0·665	−0·326	0·167	1·359	−0·441	0·583	−1·006

The 501st and 502nd terms are 0·663 and 0·452 respectively.

Table 18. *Series* 16

No. of term	0+	50+	100+	150+	200+	250+	300+	350+	400+	450+
1	3·064	2·563	−0·764	−0·950	−0·510	0·782	1·212	2·643	−4·404	−1·514
2	3·961	2·048	−1·840	−0·842	−0·417	−0·229	1·060	0·389	−4·962	−0·623
3	4·084	0·368	−2·607	−0·439	0·382	1·479	1·022	−3·470	−2·520	−0·478
4	2·641	−1·487	−3·376	−0·353	1·305	2·379	−0·085	−5·591	0·545	0·548
5	0·387	−3·636	−3·035	0·864	0·654	1·475	0·732	−4·392	2·663	−0·112
6	−1·705	−2·287	−3·348	2·270	0·538	0·155	1·083	−2·209	3·704	1·287
7	−2·375	0·076	−0·134	3·301	0·933	−2·173	2·033	−0·352	3·730	3·293
8	−1·587	1·768	0·656	2·950	0·807	−3·615	1·701	−0·099	1·325	2·896
9	−1·359	2·401	0·885	1·698	1·007	−2·885	−0·361	1·063	−0·677	0·508
10	−0·922	1·415	−0·558	0·345	0·869	−1·085	−2·497	1·611	−0·981	−1·852
11	−1·952	0·306	−0·147	−1·101	1·569	2·399	−4·262	0·638	−0·995	−0·276
12	−2·974	−0·464	0·293	−2·360	1·479	2·024	−2·711	−0·333	−1·763	−0·283
13	−2·327	0·371	1·829	−1·439	0·424	0·329	−2·430	−1·897	0·662	−1·242
14	−1·644	0·886	0·592	1·311	0·521	−1·161	−0·075	−1·415	1·912	−1·111
15	1·171	0·596	−2·140	1·653	0·935	−1·954	2·836	−1·942	2·794	−1·123
16	2·452	0·720	−2·421	2·216	1·745	−1·682	3·317	−1·781	2·191	−0·742
17	2·732	0·819	−1·626	2·804	2·092	−1·674	1·154	−0·137	1·028	−1·375
18	2·185	0·501	−0·304	4·298	1·670	−1·200	−1·548	1·512	0·237	0·075
19	−0·051	0·452	1·202	3·422	0·189	−1·420	−3·281	1·651	0·149	2·419
20	−1·314	−0·067	1·317	2·071	−0·102	−2·066	−2·276	1·924	0·228	2·651
21	−1·625	0·441	0·576	−1·412	1·012	−0·536	−0·709	1·975	0·780	0·746
22	−1·548	2·645	−0·150	−1·968	1·224	0·157	0·471	0·848	−0·034	−0·740
23	−1·434	1·854	1·152	−0·102	0·828	−1·633	2·009	1·009	0·540	−1·622
24	−0·152	−1·178	1·224	1·268	−0·221	−1·693	2·228	−0·410	1·793	−1·534
25	1·268	−2·186	0·286	1·078	0·727	−0·586	1·674	−1·272	1·069	0·325
26	0·311	−1·884	−0·847	−0·758	2·894	0·573	1·440	−2·343	−1·307	1·269
27	0·104	0·353	−1·311	−1·662	1·397	0·524	0·066	−1·628	−3·096	0·985
28	0·353	1·487	−1·573	−1·552	−1·280	0·685	−0·813	−1·354	−2·303	1·057
29	0·796	1·278	−0·506	−0·472	−2·184	−0·447	−1·714	0·739	−0·421	−0·121
30	1·040	0·392	1·700	2·173	−0·730	−0·708	−1·864	0·406	1·364	1·489
31	0·197	−0·944	3·175	1·939	1·304	0·181	−0·467	−1·072	1·436	1·055
32	−0·839	−2·411	2·770	−0·975	3·031	0·309	1·474	−2·312	2·206	1·470
33	−2·624	−0·517	2·848	−2·957	1·874	−0·529	1·485	−1·450	1·130	2·937
34	−2·509	2·182	1·846	−3·524	1·190	−1·019	2·173	−0·260	−0·794	3·681
35	−1·461	2·334	0·963	−2·963	2·232	−0·729	1·838	1·457	−2·513	0·759
36	−0·403	1·409	−1·119	−2·875	3·858	−0·223	0·463	0·872	−3·121	−1·468
37	1·415	0·978	−1·409	−1·629	2·959	−2·154	0·051	−1·933	−1·857	−2·042
38	3·393	0·876	0·453	0·828	2·638	−2·989	0·607	−1·806	−0·964	−1·349
39	1·468	−0·164	3·249	3·574	2·078	−4·117	−0·063	−1·667	−1·240	−0·791
40	1·314	−0·206	4·693	3·877	2·361	−3·747	0·854	−0·199	−2·515	−0·092
41	2·495	−0·784	4·654	2·884	1·805	−2·668	2·892	2·062	−2·841	−0·795
42	3·212	−1·908	1·653	−0·167	1·509	−0·920	2·966	3·637	−1·908	0·040
43	3·515	−1·189	−0·513	−2·624	0·535	0·748	2·061	2·915	−0·707	0·960
44	2·043	−0·502	−0·862	−3·502	1·200	−1·446	−0·720	2·352	−0·444	0·122
45	0·941	−0·066	0·439	−3·353	0·127	−3·917	−1·453	0·757	−0·304	−0·292
46	1·412	−0·840	1·859	−2·223	−3·203	−3·804	−2·092	−1·551	−0·119	−1·741
47	1·412	−1·547	2·675	−0·853	−3·135	−1·377	−1·266	−1·182	0·477	−1·774
48	0·345	−1·264	2·299	0·838	−2·173	0·555	1·012	−0·966	1·167	−2·087
49	1·749	−2·379	0·647	0·863	−0·912	2·062	1·347	−1·504	1·277	−0·746
50	2·864	−1·177	−0·412	0·544	0·069	2·761	1·904	−1·732	−1·009	0·675

Table 19. *Values of c_k and r_k for series 1*

k	c_k	r_k	k	c_k	r_k	k	c_k	r_k
0	1,217,176	1·000	17	76,902	0·066	34	54,121	0·048
1	927,053	0·763	18	37,770	0·032	35	9,302	0·008
2	458,107	0·378	19	−8,169	−0·007	36	−41,850	−0·037
3	96,479	0·080	20	−29,750	−0·026	37	−52,745	−0·047
4	−80,933	−0·067	21	−6,035	−0·005	38	−26,534	−0·024
5	−95,200	−0·079	22	22,188	0·019	39	−19,038	−0·017
6	−47,312	−0·039	23	12,361	0·011	40	−33,765	−0·030
7	−8,895	−0·007	24	−12,771	−0·011	41	−43,076	−0·039
8	26,685	0·022	25	−47,794	−0·041	42	−13,075	−0·012
9	21,158	0·018	26	−91,940	−0·080	43	37,682	0·034
10	−42,728	−0·036	27	−125,573	−0·109	44	61,317	0·055
11	−123,352	−0·104	28	−146,002	−0·127	45	53,467	0·048
12	−173,350	−0·146	29	−134,177	−0·117	46	44,736	0·041
13	−153,577	−0·130	30	−86,031	−0·075	47	45,566	0·041
14	−62,428	−0·053	31	−20,509	−0·018	48	38,801	0·035
15	34,670	0·029	32	44,838	0·039	49	24,435	0·022
16	82,284	0·070	33	77,980	0·069	50	7,831	0·007

Table 20. *Values of c_k and r_k for series 2, 3 and 4*

	Series 2		Series 3		Series 4	
k	c_k	r_k	c_k	r_k	c_k	r_k
0	835,517	1·000	939,563	1·000	480,312	1·000
1	709,079	0·852	580,821	0·621	−327,751	−0·685
2	505,325	0·610	−97,646	−0·105	83,299	0·175
3	291,381	0·353	−563,764	−0·608	124,239	0·262
4	104,876	0·128	−549,785	−0·595	−195,590	−0·414
5	−36,275	−0·044	−164,843	−0·179	142,973	0·304
6	−102,830	−0·126	235,473	0·257	−59,201	−0·126
7	−121,945	−0·150	362,975	0·398	−3,399	−0·007
8	−114,325	−0·142	221,837	0·244	7,358	0·016
9	−76,997	−0·096	−10,213	−0·011	29,175	0·063
10	−46,795	−0·058	−178,603	−0·198	−54,608	−0·119
11	−26,232	−0·033	−200,113	−0·223	48,217	0·105
12	−11,782	−0·015	−119,942	−0·134	−16,371	−0·036
13	−7,589	−0·010	−8,147	−0·009	−8,168	−0·018
14	−32,356	−0·041	111,156	0·126	−2,633	−0·006
15	−61,164	−0·078	167,156	0·190	19,762	0·044
16	−102,551	−0·132	140,275	0·160	−21,654	−0·048
17	−131,550	−0·169	48,177	0·055	−2,216	−0·005
18	−136,460	−0·177	−59,858	−0·069	30,250	0·068
19	−120,194	−0·156	−125,095	−0·145	−41,995	−0·095
20	−78,730	−0·103	−88,301	−0·103	26,538	0·060
21	−23,001	−0·030	21,297	0·025	6,283	0·014
22	39 088	0·050	122,829	0·144	−28,023	−0·064
23	76,008	0·099	157,567	0·185	46,618	0·107
24	108,169	0·144	81,340	0·096	−51,382	−0·119
25	141,980	0·190	−63,776	−0·076	22,938	0·053
26	137,726	0·185	−171,262	−0·204	42,239	0·099
27	99,392	0·134	−168,798	−0·202	−90,303	−0·212
28	72,398	0·098	−66,814	−0·081	90,937	0·214
29	71,157	0·097	63,524	0·077	−33,782	−0·080
30	68,511	0·094	147,371	0·179	−25,722	−0·061

Table 21. *Values of c_k and r_k for series 5 and 6*

Series ...	5*a*	5*b*	5*c*	5*d*	5 (*in toto*)
c_0	977,837	958,436	879,736	872,844	3,688,853
c_1	739,993	702,480	607,461	619,552	2,672,905
r_1	0·7568	0·7329	0·6905	0·7098	0·7246
c_2	364,400	282,372	168,243	223,038	1,043,493
r_2	0·3727	0·2946	0·1912	0·2555	0·2829
c_3	58,786	−60,840	−128,204	−98,570	−221,255
r_3	0·0601	−0·0635	−0·1457	−0·1129	−0·0600
c_4	−90,575	−243,083	−198,042	−276,361	−804,161
r_4	−0·0926	−0·2536	−0·2251	−0·3166	−0·2180
Series ...	6*a*	6*b*	6*c*	6*d*	6 (*in toto*)
c_0	1,551,606	1,505,293	1,548,545	1,186,705	5,792,149
c_1	924,569	903,559	951,497	750,232	3,531,836
r_1	0·5959	0·6003	0·6144	0·6322	0·6098
c_2	−270,379	−220,471	−194,686	−48,120	−720,610
r_2	−0·1743	−0·1465	−0·1257	−0·0405	−0·1244
c_3	−1,021,112	−927,370	−976,079	−604,656	−3,509,827
r_3	−0·6581	−0·6161	−0·6303	−0·5095	−0·6060
c_4	−872,747	−793,867	−901,618	−645,513	−3,206,192
r_4	−0·5625	−0·5274	−0·5822	−0·5440	−0·5535

Table 22. *Values of c_k and r_k for series 8, 10, 12, 14 and 16*

k	Series 8		Series 10		Series 12		Series 14		Series 16	
	c_k	r_k	c_k	r_k	c_k	r_k	c_k	r_k	c_k	r_k
0	12,406·47	1·000	6414·470	1·000	3346·111	1·000	2115·829	1·000	1672·212	1·000
1	11,751·63	0·949	5707·452	0·892	2815·257	0·843	1707·930	0·809	1253·272	0·751
2	10,182·50	0·824	4131·158	0·647	1714·249	0·514	973·957	0·462	489·683	0·294
3	8,303·08	0·673	2407·876	0·378	650·804	0·196	358·001	0·170	−133·027	−0·080
4	6,590·10	0·535	1064·645	0·167	−49·417	−0·015	2·327	0·001	−387·373	−0·234
5	5,299·99	0·432	302·223	0·048	−408·133	−0·123	−132·306	−0·063	−311·683	−0·188
6	4,410·17	0·360	39·600	0·006	−548·680	−0·166	−143·107	−0·068	−85·278	−0·052
7	3,865·31	0·316	90·769	0·014	−532·595	−0·161	−143·734	−0·069	125·455	0·076
8	3,502·73	0·287	274·386	0·043	−395·760	−0·120	−144·603	−0·069	245·460	0·149
9	3,211·80	0·264	471·979	0·075	−188·798	−0·057	−116·614	−0·056	271·914	0·166
10	2,959·40	0·243	624·129	0·099	36·728	0·011	−49·845	−0·024	197·670	0·121
11	2,767·02	0·228	725·064	0·116	189·907	0·058	13·060	0·006	80·693	0·049
12	2,660·60	0·220	771·678	0·123	231·281	0·071	28·253	0·014	−11·174	−0·007
13	2,627·86	0·217	768·202	0·123	182·206	0·056	−4·922	−0·002	−49·764	−0·031
14	2·632·93	0·218	682·319	0·109	150·430	0·046	−61·316	−0·030	−30·803	−0·019
15	2,639·50	0·219	550·040	0·088	175·750	0·054	−119·775	−0·058	−8·137	−0·005
16	2,632·87	0·219	396·900	0·064	196·577	0·061	−147·510	−0·072	−9·094	−0·006
17	2,570·85	0·215	221·944	0·036	192·876	0·060	−148·196	−0·073	−26·163	−0·016
18	2,425·49	0·203	67·235	0·011	190·296	0·059	−102·048	−0·050	−53·165	−0·033
19	2,201·66	0·184	−40·242	−0·007	219·116	0·068	−21·133	−0·010	−54·609	−0·034
20	1,924·49	0·162	−131·413	−0·021	308·266	0·096	60·353	0·030	−23·027	−0·014
21	1,622·42	0·137	−216·876	−0·035	416·661	0·130	141·538	0·070	69·210	0·043
22	1,300·73	0·110	−294·638	−0·048	497·343	0·155	181·089	0·090	169·082	0·106
23	957·63	0·081	−354·742	−0·058	501·373	0·157	159·227	0·079	225·251	0·141
24	622·43	0·053	−416·650	−0·068	453·351	0·142	114·051	0·057	229·216	0·144
25	347·31	0·029	−465·492	−0·076	357·632	0·113	52·435	0·026	194·919	0·123
26	150·48	0·013	−430·809	−0·071	199·598	0·063	−13·199	−0·007	140·516	0·089
27	23·52	0·002	−272·333	−0·045	−24·634	−0·008	−26·853	−0·013	51·111	0·032
28	−63·54	−0·005	−6·663	−0·001	−221·975	−0·070	17·009	0·009	−88·074	−0·056
29	−147·03	−0·013	282·596	0·047	−292·376	−0·093	89·160	0·045	−240·431	−0·153
30	−241·18	−0·021	503·641	0·084	−248·584	−0·079	164·033	0·082	−330·713	−0·210

TABLES FOR USE IN COMPARISONS WHOSE ACCURACY INVOLVES TWO VARIANCES, SEPARATELY ESTIMATED

By ALICE A. ASPIN

Introduction

The tables are designed for use when the precision of an estimate, y, of a population parameter, η, depends linearly on two population variances, σ_1^2 and σ_2^2, the sampling variance of y being therefore of the form $(\lambda_1\sigma_1^2+\lambda_2\sigma_2^2)$, where λ_1 and λ_2 are known positive constants. If s_1^2 and s_2^2 are estimates of σ_1^2 and σ_2^2, based on f_1 and f_2 degrees of freedom, respectively, then the tables give, for two probability levels, critical values of the ratio

$$v = (y-\eta)/\sqrt{(\lambda_1 s_1^2+\lambda_2 s_2^2)}.$$

The tables are based on the assumptions of normal theory, the methods used in their calculation having been described by B. L. Welch (1947) and A. A. Aspin (1948). In particular, it is assumed that s_1^2 and s_2^2 are distributed independently of each other and of y.

Even under the restrictive assumptions of normal theory, it is not possible to table critical values of v which are fixed independently of all the sample statistics. For a given probability level, ϵ, these critical values depend on f_1 and f_2 and also on the ratio

$$c = \lambda_1 s_1^2/(\lambda_1 s_1^2+\lambda_2 s_2^2),$$

i.e. on the relative magnitudes of the observed sample variances. The function to be tabled, therefore, involves four arguments and may be written

$$v\left(f_1, f_2, \frac{\lambda_1 s_1^2}{\lambda_1 s_1^2+\lambda_2 s_2^2}, \epsilon\right).$$

It has the property that the probability,

$$\Pr.\left[\frac{(y-\eta)}{\sqrt{(\lambda_1 s_1^2+\lambda_2 s_2^2)}} > v\left(f_1, f_2, \frac{\lambda_1 s_1^2}{\lambda_1 s_1^2+\lambda_2 s_2^2}, \epsilon\right)\right], \tag{1}$$

equals ϵ, whatever may be the values of the unknown population variances, σ_1^2 and σ_2^2.

Two probability levels ($\epsilon = 0{\cdot}05$ and $\epsilon = 0{\cdot}01$) are tabled on separate pages, the object in each case being to cover, in a small compass, a wide range of the other arguments involved. Direct linear interpolation should be used everywhere except, for each f, in the panel which includes $f = \infty$. Here harmonic linear interpolation will be needed.

Some remarks on the accuracy of the tables and of certain approximations, which may be used for other probability levels, are made later in an Appendix by B. L. Welch (in particular see § 6).

Example

If $(\alpha_1-\alpha_2)$ is the difference between the means of two normal populations whose standard deviations cannot reasonably be assumed equal; and if $(\bar{x}_1-\bar{x}_2)$ is the difference between the two means of samples of sizes n_1 and n_2 taken respectively from the populations; then we may put

$$v = \frac{(\bar{x}_1-\bar{x}_2)-(\alpha_1-\alpha_2)}{\sqrt{(s_1^2/n_1+s_2^2/n_2)}}, \quad \frac{\lambda_1 s_1^2}{(\lambda_1 s_1^2+\lambda_2 s_2^2)} = \frac{s_1^2/n_1}{(s_1^2/n_1+s_2^2/n_2)},$$

$$f_1 = (n_1-1) \quad \text{and} \quad f_2 = (n_2-1),$$

Table 1. *Value of* $v = \dfrac{(y-\eta)}{\sqrt{(\lambda_1 s_1^2 + \lambda_2 s_2^2)}}$ *exceeded with probability* $\epsilon = 0{\cdot}05$*

[or of $|v|$ *exceeded with probability* $2\epsilon = 0{\cdot}10$*]*

	$\frac{\lambda_1 s_1^2}{\lambda_1 s_1^2 + \lambda_2 s_2^2}$	0·0	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
$f_2 = 6$	$f_1 = 6$	1·94	1·90	1·85	1·80	1·76	1·74	1·76	1·80	1·85	1·90	1·94
	8	1·94	1·90	1·85	1·80	1·76	1·73	1·74	1·76	1·79	1·82	1·86
	10	1·94	1·90	1·85	1·80	1·76	1·73	1·73	1·74	1·76	1·78	1·81
	15	1·94	1·90	1·85	1·80	1·76	1·73	1·71	1·71	1·72	1·73	1·75
	20	1·94	1·90	1·85	1·80	1·76	1·73	1·71	1·70	1·70	1·71	1·72
	∞	1·94	1·90	1·85	1·80	1·76	1·72	1·69	1·67	1·66	1·65	1·64
$f_2 = 8$	$f_1 = 6$	1·86	1·82	1·79	1·76	1·74	1·73	1·76	1·80	1·85	1·90	1·94
	8	1·86	1·82	1·79	1·76	1·73	1·73	1·73	1·76	1·79	1·82	1·86
	10	1·86	1·82	1·79	1·76	1·73	1·72	1·72	1·74	1·76	1·78	1·81
	15	1·86	1·82	1·79	1·76	1·73	1·71	1·71	1·71	1·72	1·73	1·75
	20	1·86	1·82	1·79	1·76	1·73	1·71	1·70	1·70	1·70	1·71	1·72
	∞	1·86	1·82	1·79	1·75	1·72	1·70	1·68	1·66	1·65	1·65	1·64
$f_2 = 10$	$f_1 = 6$	1·81	1·78	1·76	1·74	1·72	1·73	1·76	1·80	1·85	1·90	1·94
	8	1·81	1·78	1·76	1·74	1·72	1·72	1·73	1·76	1·79	1·82	1·86
	10	1·81	1·78	1·76	1·73	1·72	1·71	1·72	1·73	1·76	1·78	1·81
	15	1·81	1·78	1·76	1·73	1·72	1·70	1·70	1·71	1·72	1·73	1·75
	20	1·81	1·78	1·76	1·73	1·71	1·70	1·69	1·69	1·70	1·71	1·72
	∞	1·81	1·78	1·76	1·73	1·71	1·69	1·67	1·66	1·65	1·65	1·64
$f_2 = 15$	$f_1 = 6$	1·75	1·73	1·72	1·71	1·71	1·73	1·76	1·80	1·85	1·90	1·94
	8	1·75	1·73	1·72	1·71	1·71	1·71	1·73	1·76	1·79	1·82	1·86
	10	1·75	1·73	1·72	1·71	1·70	1·70	1·72	1·73	1·76	1·78	1·81
	15	1·75	1·73	1·72	1·70	1·70	1·69	1·70	1·70	1·72	1·73	1·75
	20	1·75	1·73	1·72	1·70	1·69	1·69	1·69	1·69	1·70	1·71	1·72
	∞	1·75	1·73	1·72	1·70	1·68	1·67	1·66	1·65	1·65	1·65	1·64
$f_2 = 20$	$f_1 = 6$	1·72	1·71	1·70	1·70	1·71	1·73	1·76	1·80	1·85	1·90	1·94
	8	1·72	1·71	1·70	1·70	1·70	1·71	1·73	1·76	1·79	1·82	1·86
	10	1·72	1·71	1·70	1·69	1·69	1·70	1·71	1·73	1·76	1·78	1·81
	15	1·72	1·71	1·70	1·69	1·69	1·69	1·69	1·70	1·72	1·73	1·75
	20	1·72	1·71	1·70	1·69	1·68	1·68	1·68	1·69	1·70	1·71	1·72
	∞	1·72	1·71	1·70	1·68	1·67	1·66	1·66	1·65	1·65	1·65	1·64
$f_2 = \infty$	$f_1 = 6$	1·64	1·65	1·66	1·67	1·69	1·72	1·76	1·80	1·85	1·90	1·94
	8	1·64	1·65	1·65	1·66	1·68	1·70	1·72	1·75	1·79	1·82	1·86
	10	1·64	1·65	1·65	1·66	1·67	1·69	1·71	1·73	1·76	1·78	1·81
	15	1·64	1·65	1·65	1·65	1·66	1·67	1·68	1·70	1·72	1·73	1·75
	20	1·64	1·65	1·65	1·65	1·66	1·66	1·67	1·68	1·70	1·71	1·72
	∞	1·64	1·64	1·64	1·64	1·64	1·64	1·64	1·64	1·64	1·64	1·64

* y is normally distributed about η with variance $\sqrt{(\lambda_1 \sigma_1^2 + \lambda_2 \sigma_2^2)}$, and s_1^2 and s_2^2 are independent estimates of σ_1^2 and σ_2^2, based on f_1 and f_2 degrees of freedom, respectively. λ_1 and λ_2 are known constants.

In the problem of comparing the means of samples taken from two normal populations, put $y = (\bar{x}_1 - \bar{x}_2)$, $f_1 = (n_1 - 1)$, $f_2 = (n_2 - 1)$, $\lambda_1 = 1/n_1$ and $\lambda_2 = 1/n_2$, where n_1 and n_2 are the sample sizes.

Table 2. *Value of* $v = \dfrac{(y-\eta)}{\sqrt{(\lambda_1 s_1^2 + \lambda_2 s_2^2)}}$ *exceeded with probability* $\epsilon = 0{\cdot}01$*

[or of $|v|$ *exceeded with probability* $2\epsilon = 0{\cdot}02$*]*

	$\frac{\lambda_1 s_1^2}{\lambda_1 s_1^2 + \lambda_2 s_2^2}$	0·0	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
$f_2 = 10$	$f_1 = 10$	2·76	2·70	2·63	2·56	2·51	2·50	2·51	2·56	2·63	2·70	2·76
	12	2·76	2·70	2·63	2·56	2·51	2·49	2·49	2·52	2·57	2·62	2·68
	15	2·76	2·70	2·63	2·56	2·51	2·48	2·47	2·48	2·52	2·56	2·60
	20	2·76	2·70	2·63	2·56	2·51	2·47	2·45	2·45	2·47	2·49	2·53
	30	2·76	2·70	2·63	2·56	2·50	2·46	2·43	2·42	2·42	2·44	2·46
	∞	2·76	2·70	2·63	2·56	2·50	2·44	2·40	2·36	2·34	2·33	2·33
$f_2 = 12$	$f_1 = 10$	2·68	2·62	2·57	2·52	2·49	2·49	2·51	2·56	2·63	2·70	2·76
	12	2·68	2·62	2·57	2·52	2·48	2·47	2·48	2·52	2·57	2·62	2·68
	15	2·68	2·62	2·57	2·52	2·48	2·46	2·46	2·48	2·52	2·56	2·60
	20	2·68	2·62	2·57	2·52	2·48	2·45	2·44	2·45	2·47	2·49	2·53
	30	2·68	2·62	2·57	2·52	2·47	2·44	2·42	2·41	2·42	2·44	2·46
	∞	2·68	2·62	2·57	2·51	2·46	2·42	2·38	2·36	2·34	2·33	2·33
$f_2 = 15$	$f_1 = 10$	2·60	2·56	2·52	2·48	2·47	2·48	2·51	2·56	2·63	2·70	2·76
	12	2·60	2·56	2·52	2·48	2·46	2·46	2·48	2·52	2·57	2·62	2·68
	15	2·60	2·56	2·51	2·48	2·45	2·45	2·45	2·48	2·51	2·56	2·60
	20	2·60	2·56	2·51	2·48	2·45	2·43	2·43	2·44	2·46	2·49	2·53
	30	2·60	2·56	2·51	2·47	2·44	2·42	2·41	2·41	2·42	2·44	2·46
	∞	2·60	2·56	2·51	2·47	2·43	2·40	2·37	2·35	2·34	2·33	2·33
$f_2 = 20$	$f_1 = 10$	2·53	2·49	2·47	2·45	2·45	2·47	2·51	2·56	2·63	2·70	2·76
	12	2·53	2·49	2·47	2·45	2·44	2·45	2·48	2·52	2·57	2·62	2·68
	15	2·53	2·49	2·46	2·44	2·43	2·43	2·45	2·48	2·51	2·56	2·60
	20	2·53	2·49	2·46	2·44	2·42	2·42	2·42	2·44	2·46	2·49	2·53
	30	2·53	2·49	2·46	2·44	2·42	2·40	2·40	2·40	2·42	2·43	2·46
	∞	2·53	2·49	2·46	2·43	2·40	2·38	2·36	2·34	2·33	2·33	2·33
$f_2 = 30$	$f_1 = 10$	2·46	2·44	2·42	2·42	2·43	2·46	2·50	2·56	2·63	2·70	2·76
	12	2·46	2·44	2·42	2·41	2·42	2·44	2·47	2·52	2·57	2·62	2·68
	15	2·46	2·44	2·42	2·41	2·41	2·42	2·44	2·47	2·51	2·56	2·60
	20	2·46	2·43	2·42	2·40	2·40	2·40	2·42	2·44	2·46	2·49	2·53
	30	2·46	2·43	2·42	2·40	2·39	2·39	2·39	2·40	2·42	2·43	2·46
	∞	2·46	2·43	2·41	2·39	2·37	2·36	2·35	2·34	2·33	2·33	2·33
$f_2 = \infty$	$f_1 = 10$	2·33	2·33	2·34	2·36	2·40	2·44	2·50	2·56	2·63	2·70	2·76
	12	2·33	2·33	2·34	2·36	2·38	2·42	2·46	2·51	2·57	2·62	2·68
	15	2·33	2·33	2·34	2·35	2·37	2·40	2·43	2·47	2·51	2·56	2·60
	20	2·33	2·33	2·33	2·34	2·36	2·38	2·40	2·43	2·46	2·49	2·53
	30	2·33	2·33	2·33	2·34	2·35	2·36	2·37	2·39	2·41	2·43	2·46
	∞	2·33	2·33	2·33	2·33	2·33	2·33	2·33	2·33	2·33	2·33	2·33

* y is normally distributed about η with variance $\sqrt{(\lambda_1 \sigma_1^2 + \lambda_2 \sigma_2^2)}$, and s_1^2 and s_2^2 are independent estimates of σ_1^2 and σ_2^2, based on f_1 and f_2 degrees of freedom, respectively. λ_1 and λ_2 are known constants.

In the problem of comparing the means of samples taken from two normal populations, put $y = (\bar{x}_1 - \bar{x}_2)$, $f_1 = (n_1 - 1)$, $f_2 = (n_2 - 1)$, $\lambda_1 = 1/n_1$ and $\lambda_2 = 1/n_2$, where n_1 and n_2 are the sample sizes.

where s_1^2 and s_2^2 are the sample estimates of variance. The tables may then be used to make inferences about $(\alpha_1 - \alpha_2)$.

Thus if $n_1 = 10$, $n_2 = 15$, $\bar{x}_1 = 73{\cdot}4$, $\bar{x}_2 = 47{\cdot}1$, $s_1^2 = 51$, and $s_2^2 = 141$, we shall have $(\bar{x}_1 - \bar{x}_2) = 26{\cdot}3, f_1 = 9, f_2 = 14$ and

$$v = \frac{26{\cdot}3 - (\alpha_1 - \alpha_2)}{3{\cdot}81}, \quad \frac{\lambda_1 s_1^2}{(\lambda_1 s_1^2 + \lambda_2 s_2^2)} = 0{\cdot}352.$$

From the tables $v\left\{f_1, f_2, \frac{\lambda_1 s_1^2}{\lambda_1 s_1^2 + \lambda_2 s_2^2}, 0{\cdot}05\right\} = v\{9, 14, 0{\cdot}352, 0{\cdot}05\} = 1{\cdot}71.$

If it were a matter of testing the consistency of the data with the hypothesis that $\alpha_1 = \alpha_2$, we should have $v = 26{\cdot}3/3{\cdot}81 = 6{\cdot}90$, which is clearly far beyond the tabled 5 % point.

In obtaining an interval estimate for $(\alpha_1 - \alpha_2)$ we should note that in the above description of the tables we have been dealing explicitly with one-sided probabilities. The chance that v exceeds the tabled value *numerically*, either in the positive or negative direction, is 2ϵ. Thus

$$\Pr.\left[-1{\cdot}71 < \frac{26{\cdot}3 - (\alpha_1 - \alpha_2)}{3{\cdot}81} < 1{\cdot}71\right] = 0{\cdot}90,$$

i.e.
$$\Pr.[19{\cdot}8 < (\alpha_1 - \alpha_2) < 32{\cdot}8] = 0{\cdot}90,$$

i.e. the 90 % confidence limits for $(\alpha_1 - \alpha_2)$ are 19·8 and 32·8.

APPENDIX

FURTHER NOTE ON MRS ASPIN'S TABLES AND ON CERTAIN APPROXIMATIONS TO THE TABLED FUNCTION

BY B. L. WELCH

1. More than one solution has been proposed for the problem of dealing with comparisons whose estimated precision involves two separate estimates of variance. The present tables are based on a solution which relies only on probability calculations of the so-called 'direct' variety. I have given elsewhere a method of developing this solution in a power series in $1/f_i$ (Welch, 1947), and further terms in the series have been given by A. A. Aspin (1948). In calculating the tables, Mrs Aspin has in general utilized the series as far as terms of order $1/f_i^3$, and, in certain doubtful cases, has also evaluated terms of order $1/f_i^4$. Taking the series thus far, it is possible to give v to two decimal places down to the lowest values of f_1 and f_2 shown in the tables, but, for smaller f_1 and f_2, two-decimal accuracy could not, in general, have been guaranteed. For the larger values of f_1 and f_2 more figures could, of course, have been given, but there are advantages, in a table of this character, in keeping to the same number of figures throughout.

It is natural to ask what degree of accuracy in the probability is implied by the use of values of v rounded off to two decimal places. For f_1 and f_2 large, this question can easily be answered. For then v tends to be normally distributed, and, on the normal curve, two-decimal accuracy in the deviate, in the vicinity of $\epsilon = 0{\cdot}05$, implies three-decimal accuracy in the probability of the deviate being exceeded. In the vicinity of $\epsilon = 0{\cdot}01$, two-decimal accuracy in the deviate implies errors in the probability of not more than 2 units in the fourth decimal place.

For lower values of f_1 and f_2 it is not possible to be dogmatic, although it is very often the case, when we are using series of the present type, that the greater the struggle needed to obtain extra decimal places in the deviate, the less will be the effect upon the probability itself of any rounding-off error in the deviate. Although I believe this to be true in the present instance it has seemed to me, nevertheless, advisable to make some direct calculations, using quadratures, to test out the question thoroughly. Accordingly for the lowest values $f_1 = f_2 = 6$ given in the table for $\epsilon = 0{\cdot}05$, I have made some calculations of this nature and will present the results below. It will be convenient at the same time, for the same particular case, to give the results of similar calculations which throw light on the accuracy of certain simple approximations to the tabled function.

2. If we are confining ourselves to a given pair of values, f_1 and f_2, and to a given probability level, ϵ, the function $v\left\{f_1, f_2, \frac{\lambda_1 s_1^2}{\lambda_1 s_1^2 + \lambda_2 s_2^2}, \epsilon\right\}$ need, momentarily at least, be regarded only as a function of the one argument $c = \lambda_1 s_1^2/(\lambda_1 s_1^2 + \lambda_2 s_2^2)$ and may conveniently be denoted simply by $v(c)$. Since it is not the function $v(c)$ itself that we are concerned with, but the tabular values obtained after rounding off $v(c)$ to two decimal places, it will be convenient to have a further symbol $v_T(c)$, say, to denote these rounded-off values. In the first line of Table A we reproduce from Mrs Aspin's Table 1, the values $v_T(c)$ for the case $f_1 = f_2 = 6$, $\epsilon = 0{\cdot}05$. Let us now consider the probability that v exceeds $v_T(c)$.

Table A. *Critical values of* v ($f_1 = f_2 = 6$; $\epsilon = 0{\cdot}05$)

$c =$	0·0 or 1·0	0·1 or 0·9	0·2 or 0·8	0·3 or 0·7	0·4 or 0·6	0·5
Tabled values $v_T(c)$	1·94	1·90	1·85	1·80	1·76	1·74
Approximation $v_1(c)$	1·94	1·88	1·84	1·81	1·79	1·78
Approximation $v_1'(c)$	1·94	1·87	1·82	1·79	1·77	1·76

If Pr. $\{v > v_T(c)\}$ depends on σ_1^2 and σ_2^2 at all, it must depend on them through the ratio $\gamma = \lambda_1\sigma_1^2/(\lambda_1\sigma_1^2 + \lambda_2\sigma_2^2)$. If, then, we take different values of γ, and calculate for each one, directly by *quadrature*, the value Pr. $\{v > v_T(c)\}$, we shall have solved our problem. This calculation by quadrature is heavy, although it is lightened considerably by the use of some rather neglected tables of the probability integral of 'Student's' distribution (Gosset, 1925). It is not proposed to enter into the details of the calculations, but it should be mentioned that a rather broad network was used in the quadrature and that the last decimal place given below cannot, on this account, be fully guaranteed.

The following results were obtained. For $\gamma = 0{\cdot}1$, $0{\cdot}2$, $0{\cdot}3$, $0{\cdot}4$ and $0{\cdot}5$ the respective values for Pr. $\{v > v_T(c)\}$ were 0·0501, 0·0500, 0·0500, 0·0498 and 0·0498. It is clear, therefore, that to three decimals we certainly have Pr. $\{v > v_T(c)\} = 0{\cdot}050$, whatever γ, and indeed that the error can, at most, be only a unit or two in the fourth decimal place. It seems, therefore, safe to assume that *throughout* Table 1 the rounding off to two decimal places in the deviate leaves us easily with three decimal places in the probability, 0·050, that the deviate is exceeded. Similarly in Table 2 we can almost everywhere expect that the probability that v exceeds the tabled values will be 0·0100 to four decimal places.

3. The provision of tables of v covering a large number of probability levels would clearly be a major task, and possibly one which would not justify the labour involved. It is therefore of some importance to consider the order of accuracy attainable with any approximation which only utilizes existing tables. A simple approximation of this character is to take for the critical v the value given in a 'Student' t table corresponding to degrees of freedom F, where

$$1/F = \{c^2/f_1 + (1-c)^2/f_2\}.$$

Consideration of the series development of the 'Student' deviate shows that this procedure is legitimate if terms of order $1/f_i^2$ can be neglected.

Numerically, in the particular case $f_1 = f_2 = 6$ with $\epsilon = 0{\cdot}05$, this approximation is equivalent to taking for $v(c)$ the values $v_1(c)$ shown in the second line of Table A. By direct evaluation, using quadratures, it is possible now to obtain the probability that $v > v_1(c)$. Results of this calculation, for different γ, are shown in the second line of Table B. In this case an error

Table B

$\gamma=$	0·0 or 1·0	0·1 or 0·9	0·2 or 0·8	0·3 or 0·7	0·4 or 0·6	0·5
Pr. $\{v > v_F(c)\}$	0·050	0·050	0·050	0·050	0·050	0·050
Pr. $\{v > v_1(c)\}$	0·050	0·051	0·050	0·049	0·048	0·048
Pr. $\{v > v_1'(c)\}$	0·050	0·052	0·051	0·051	0·050	0·050

of one or two units in the *third* decimal place of the probability is apparently possible with this approximation.

As f_1 and f_2 increase, the approximation using the 'Student' table improves, and for practical purposes this procedure is therefore extremely useful. The approximation can also be used in conjunction with the above-mentioned tables of the probability integral of the 'Student' distribution (Gosset, 1925) to give actual probabilities, if it is felt that reference to a few percentage points is not sufficient.

4. To order $1/f_i$, the above-described procedure will not be altered if we enter the 'Student' t table with some other number of degrees of freedom, F' (say), which differs from F only to order $1/f_i^2$. A possible modification (see Welch, 1947, p. 32, equation (29)) is to take F' given by

$$1/(F'+2) = \{c^2/(f_1+2) + (1-c)^2/(f_2+2)\}.$$

A form like this is suggested by the behaviour of the successive groups of terms in the true series solution. An expansion of this series in terms of $1/(f_i+2)$, $1/(f_i+2)(f_i+4)$, etc., instead of powers of $1/f_i$ appears to introduce some simplification into the algebra. Numerical calculations, however, show this gain to be largely illusory (Aspin, 1948), and the corresponding 'Student' approximation, using F', not to have the advantages expected.

In the third line of Table A are shown the values $v_1'(c)$ which result when F' is used, and the third line in Table B shows values of Pr. $\{v > v_1'(c)\}$ calculated by quadrature. In the neighbourhood of $\gamma = \frac{1}{2}$, F' has some advantage, but, taking the overall picture, the accuracy given is of the same order as that given by using the F of the previous section. Since F is, perhaps, rather simpler to use, we have therefore no very good reason to introduce a modification of the present nature.

5. Approximations, utilizing tables of the 'Student' distribution and agreeing with the true solution to terms of order $1/f_i^3$ or higher, can be given. The labour involved in using them, however, begins to approach that involved in working out the true series solution term by term, and there seems, therefore, little point in describing them.

6. *Summary.* On the basis of the above calculations for the case $f_1 = f_2 = 6$ and the known behaviour of the distribution of v when f_1 and f_2 are large, it seems safe to say, that, although the values of v given in Table 1 are rounded off to two decimal places, the probability of the values thus tabled being exceeded will always equal 0·050 to three decimal places. Usually, indeed, the accuracy of the probability will be even better than this.

Similar considerations for Table 2 suggest that the probability of the given values of v being exceeded will almost everywhere equal 0·0100 to four decimal places.

The above calculations also throw some light on an approximate method which consists in taking for the required v, the values in a 'Student' t table corresponding to F degrees of freedom, where

$$1/F = \{c^2/f_1 + (1-c)^2/f_2\} \quad \text{and} \quad c = \lambda_1 s_1^2/(\lambda_1 s_1^2 + \lambda_2 s_2^2).$$

If this approximation be used for values of the arguments not covered by Mrs Aspin's tables, or in cases where actual probabilities are required rather than references to fixed probability levels, it usually gives good results.

REFERENCES

Aspin, A. A. (1948). An examination and further development of a formula occurring in the problem of comparing two mean values. *Biometrika*, **35**, 88–96.

Gosset, W. S. (1925). New tables for testing the significance of observations. *Metron*, **5**, 105. (Reprinted in '*Student's*' *Collected Papers*, Cambridge University Press, 1942.)

Welch, B. L. (1947). The generalization of 'Student's' problem when several different population variances are involved. *Biometrika*, **34**, 28–35.

BIVARIATE DISTRIBUTIONS BASED ON SIMPLE TRANSLATION SYSTEMS

By N. L. JOHNSON

1. Introduction

In a previous paper (Johnson, 1949) systems of frequency curves based on transformations of the form

$$z = \gamma + \delta f\left(\frac{x-\xi}{\lambda}\right) = \gamma + \delta f(y),$$

where z is a unit Normal variable, f is a function, and γ, δ, ξ and λ are parameters, have been discussed. The special systems

$$S_L \text{ corresponding to } f(y) = \log y \text{ ('log Normal')},$$

$$S_B \text{ corresponding to } f(y) = \log \frac{y}{1-y},$$

and

$$S_U \text{ corresponding to } f(y) = \log [y + \sqrt{(y^2+1)}]$$

were considered in some detail.

The present paper is concerned with certain bivariate distributions which can be constructed on the basis of univariate distributions of systems S_L, S_B or S_U, together with the Normal distribution, denoted below as S_N.

If it be supposed that the unit Normal variables

$$z_1 = \gamma_1 + \delta_1 f_1(y_1), \quad z_2 = \gamma_2 + \delta_2 f_2(y_2) \tag{1}$$

have the joint Normal bivariate distribution

$$p(z_1, z_2) = [2\pi\sqrt{(1-\rho^2)}]^{-1} \exp\{-\tfrac{1}{2}(1-\rho^2)^{-1}(z_1^2 - 2\rho z_1 z_2 + z_2^2)\}, \tag{2}$$

the joint distribution of y_1 and y_2 is thereby determined. While we cannot, of course, say that the joint distribution of z_1 and z_2 must be of form (2), it is an assumption which should be reasonable in many cases. With equal cogency, therefore, it can be argued that the joint distribution of the variables y_1 and y_2, defined by (1), should be that implied by (2).

The transformations for y_1 and y_2 need not belong to the same system. Thus, restricting ourselves to the Normal system, S_N, and the systems S_L, S_B and S_U we have ten different systems of bivariate distributions for y_1 and y_2. We will denote these systems by symbols S_{NN}, S_{NL}, S_{LB}, etc. S_{NN} is the bivariate Normal distribution; S_{LL} is the 'logarithmic surface' considered by Wicksell (1917); S_{NL} is the 'semi-logarithmic surface' considered by Yuan (1933). Allowing for choice of dependent and independent variable, there will be sixteen types of regression line (one being the linear regression corresponding to the bivariate Normal case).

2. Comparison with other bivariate systems

It is of interest to compare the principle employed above (transformation to Normality) with other principles which have been suggested as starting points in the construction of non-Normal bivariate distributions. Steffensen (1922, 1941) suggested that the joint distribution $p(y_1, y_2)$ might be supposed to be of the form

$$g_1(a_1 y_1 + a_2 y_2)\, g_2(b_1 y_1 + b_2 y_2),$$

where a_1, a_2, b_1, b_2 are constants and g_1, g_2, are univariate probability density functions. This is equivalent (Frechet, 1941) to supposing that independent 'primary variables' can be found which are linear functions of the y's. K. Pearson (1923), in a survey of methods of construction of bivariate distributions, states that this method does not accord well with observed distributions. In the same paper, reference is made to a number of bivariate distributions arising from special cases of a generalization of the differential equation which generates the univariate Pearson system of curves (see also Rhodes, 1922). Reference is also made to contemporaneous work by Narumi (1923), wherein bivariate distributions are derived from certain limitations on the shape of the regression lines and on the type of heteroscedasticity allowable. Pretorius (1930) showed that Narumi's assumption of constant β_1 and β_2 in the array distributions is not justifiable in much observational data. Pretorius also considered bivariate distributions based on Charlier curves, on Edgeworth's method of translation, and on a development of our system S_{LL}. Recently Van Uven (1947, 1948) has developed a complete system of surfaces generalizing the Pearson univariate system.

3. General properties of the S_{IJ} systems

(*a*) *Array distributions.* The system S_{IJ} is defined by the equations

$$z_1 = \gamma_1 + \delta_1 f_I(y_1), \quad z_2 = \gamma_2 + \delta_2 f_J(y_2) \tag{3}$$

(f_I and f_J referring to the systems S_I and S_J respectively, $I, J = N, L, B$ or U), and by the correlation, ρ, between z_1 and z_2.

If y_1 be fixed, so is z_1. For a fixed value of z_1, z_2 is distributed Normally with expected value ρz_1 and standard deviation $(1-\rho^2)^{\frac{1}{2}}$. Thus, for y_1 fixed, $\gamma_2 + \delta_2 f_J(y_2)$ is distributed Normally with expected value $\rho\{\gamma_1 + \delta_1 f_I(y_1)\}$ and standard deviation $(1-\rho^2)^{\frac{1}{2}}$. Hence

$$[\gamma_2 + \delta_2 f_J(y_2) - \rho\{\gamma_1 + \delta_1 f_I(y_1)\}](1-\rho^2)^{-\frac{1}{2}}$$

is a unit Normal variable. Thus, given y_1, y_2 has a distribution of the same system, S_J, as its marginal distribution, with

$$\gamma_2 \text{ replaced by } [\gamma_2 - \rho\{\gamma_1 + \delta_1 f_I(y_1)\}](1-\rho^2)^{-\frac{1}{2}}$$

and

$$\delta_2 \text{ replaced by } \delta_2(1-\rho^2)^{-\frac{1}{2}}.$$

Since $f_I(y_1)$ varies from $-\infty$ to $+\infty$, it is clear that (except when $\rho = 0$) the quantity replacing γ_2 will change sign at some point in the range of variation of y_1. If S_J be a symmetrical transformation then the array distributions to either side of this value of y_1 will be skew in opposite directions. Variation in skewness of this kind was present in some of the examples used by Pretorius. It may, of course, so happen that the transition point is so far out in the tail of the y_1 distribution that the reversal effect is negligible.

It may also be noted that the quantity replacing δ_2 does not vary with y_1. If the distribution of y_2 be S_B or S_U, the (β_1, β_2) points of the array distributions all fall on a constant δ line; if the distribution of y_2 be S_L, the (β_1, β_2) point does not vary, so that the array distributions are all of the same shape, though they are not homoscedastic.

A case of particular interest arises when the marginal distribution of y_2 is in the system S_B. The array distribution of y_2, given y_1, will be bimodal provided

$$\left|\frac{\gamma_2 - \rho\{\gamma_1 + \delta_1 f_I(y_1)\}}{\sqrt{(1-\rho^2)}}\right| < \frac{\sqrt{(1-\rho^2-2\delta_2^2)}}{\delta_2} - \frac{2\delta_2}{\sqrt{(1-\rho^2)}} \tanh^{-1}\sqrt{\frac{1-\rho^2-2\delta_2^2}{1-\rho^2}} \tag{4}$$

(cf. equation (26) of Johnson (1949)).

We observe that

(i) If $\rho^2 > 1 - 2\delta_2^2$, then none of the array distributions are bimodal.

(ii) There are always some unimodal array distributions. Bimodal array distributions, if any, correspond to values of y_1 in a single finite interval.

(*b*) *Median regression.* Since the array distribution of y_2, given y_1, is in the system S_J to which the marginal distribution of y_2 belongs, it follows that the expected value of y_2, given y_1, may be expressed as an explicit function of y_1 using formulae derived by Johnson (1949). This would give the regression curve of y_2 on y_1. The expression obtained would, however, in general, be complicated and unsuited to further analytical study.

Much simpler and more informative formulae are obtained if we write down the expression for the *median* value of y_2, given y_1. The possibility of using the median regression was mentioned by Narumi (1923), though he did not pursue the idea.

The median value $\tilde{y}_2$ of y_2, for a fixed value of y_1, satisfies the equation

$$\gamma_2 + \delta_2 f_J(\tilde{y}_2) = \rho[\gamma_1 + \delta_1 f_I(y_1)],$$

i.e.

$$f_J(\tilde{y}_2) = (\rho\gamma_1 - \gamma_2)\delta_2^{-1} + \rho\delta_1\delta_2^{-1} f_I(y_1). \tag{5}$$

The median regression will be linear if $\rho\delta_1 = \delta_2$, $\rho\gamma_1 = \gamma_2$ and $f_I \equiv f_J$. The only other case giving linear regression is $f_I(y) = f_J(y) = y$ which is S_{NN}, the Normal bivariate case.

Table 1 shows the median regression equations for the sixteen possible cases covered by the S_B, S_U, S_L and Normal marginal distributions. For convenience the parameters

$$\theta = e^{(\rho\gamma_1 - \gamma_2)/\delta_2}, \quad \phi = \rho\delta_1/\delta_2 \tag{6}$$

have been introduced.

Table 1

Distribution of y_2	Distribution of y_1	Median regression of y_2 on y_1 $\tilde{y}_2$
S_N	S_N	$\tilde{y}_2 = \log\theta + \phi y_1$
S_N	S_L	$\log\theta + \phi\log y_1$
S_L	S_N	$\theta e^{\phi y_1}$
S_N	S_B	$\log\theta + \phi\log\{y_1/(1-y_1)\}$
S_B	S_N	$1/(1+\theta^{-1}e^{-\phi y_1})$
S_N	S_U	$\log\theta + \phi\log[y_1 + \sqrt{(y_1^2+1)}]$
S_U	S_N	$\frac{1}{2}(\theta e^{\phi y_1} - \theta^{-1}e^{-\phi y_1})$
S_L	S_L	θy_1^{ϕ}
S_L	S_B	$\theta\{y_1/(1-y_1)\}^{\phi}$
S_B	S_L	$1/(1+\theta^{-1}y_1^{-\phi})$
S_L	S_U	$\theta[y_1 + \sqrt{(y_1^2+1)}]^{\phi}$
S_U	S_L	$\frac{1}{2}(\theta y_1^{\phi} - \theta^{-1}y_1^{-\phi})$
S_B	S_B	$\theta y_1^{\phi}/\{(1-y_1)^{\phi} + \theta y_1^{\phi}\}$
S_B	S_U	$\{1+\theta^{-1}[\sqrt{(y_1^2+1)} - y_1]^{\phi}\}^{-1}$
S_U	S_B	$\frac{1}{2}\{\theta y_1^{2\phi} - \theta^{-1}(1-y_1)^{2\phi}\}\, y_1^{-\phi}(1-y_1)^{-\phi}$
S_U	S_U	$\frac{1}{2}\{\theta[y_1 + \sqrt{(y_1^2+1)}]^{\phi} - \theta^{-1}[\sqrt{(y_1^2+1)} - y_1]^{\phi}\}$

The general appearance of these curves is indicated in Figs. 1–9 (those involving Normal marginal distributions in respect to either y_1 or y_2 are omitted). It will be noted that in many cases the general slope of the regression depends on the range of values in which ϕ lies. The case $\phi < 0$ will arise when $\rho < 0$, as we have assumed throughout this paper that δ is positive.

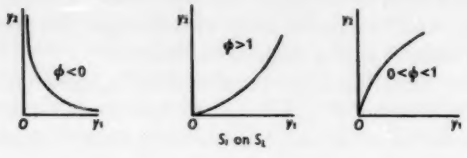

Fig. 1.

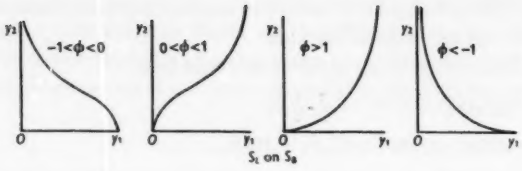

Fig. 2.

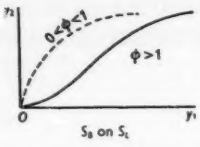

Fig. 3.

The examples in Figs. 1 and 2 should suffice to show the effect of reversing the sign of ϕ.

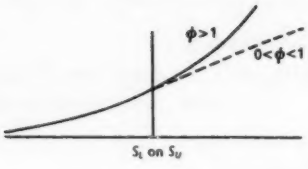

Fig. 4.

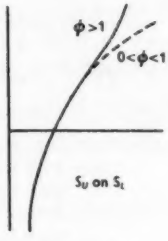

Fig. 5.

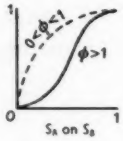

Fig. 6.

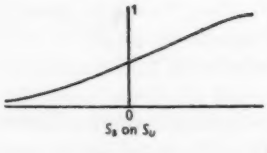

Fig. 7.

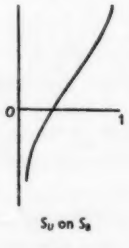

Fig. 8.

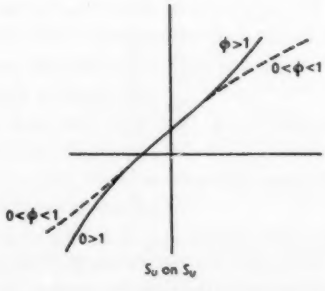

Fig. 9.

4. PERCENTAGE ZONES IN THE S_{IJ} SYSTEMS

It is, of course, a simple matter to construct regions in the (y_1, y_2) plane which will contain any desired proportion of the frequency distribution of y_1 and y_2. If k be the number defined by

$$\frac{1}{\sqrt{(2\pi)}}\int_{-k}^{k} e^{-\frac{1}{2}t^2}dt = \alpha,$$

then, for a fixed value of y_1,

$$\text{Pr.}\left\{\frac{-k-[\gamma_2-\rho\{\gamma_1+\delta_1 f_I(y_1)\}](1-\rho^2)^{-\frac{1}{2}}}{\delta_2(1-\rho^2)^{-\frac{1}{2}}} < f_J(y_2) < \frac{k-[\gamma_2-\rho\{\gamma_1+\delta_1 f_I(y_1)\}](1-\rho^2)^{-\frac{1}{2}}}{\delta_2(1-\rho^2)^{-\frac{1}{2}}}\right\} = \alpha.$$

Hence the curves

$$\pm k(1-\rho^2)^{\frac{1}{2}} + \gamma_2 + \delta_2 f_J(y_2) = \rho[\gamma_1 + \delta_1 f_I(y_1)] \qquad (7)$$

will enclose a zone containing a proportion α of the total frequency surface. This zone may be regarded as a 'percentage zone' for y_2 given y_1. It will be noted that the boundary curves (7) will be of the same type as the median regression of y_2 on y_1. They will have the same 'ϕ' ($= \rho\delta_1/\delta_2$) as the median regression, but the values of 'θ' will be

$$\exp\left[\delta_2^{-1}\{\rho\gamma_1 - \gamma_2 \pm k(1-\rho^2)^{\frac{1}{2}}\}\right].$$

Naturally there will be another similar zone containing a proportion α of the total frequency forming a 'percentage zone' for y_1 given y_2.

A third region containing a given proportion of the total frequency, which is some interest, is that obtained from the transformation of the corresponding χ^2 ellipse:

$$z_1^2 - 2\rho z_1 z_2 + z_2^2 = \text{const.} \qquad (8)$$

Substituting from (1) in (8) will give the equation of the boundary of a region in the (y_1, y_2) plane which will contain the same proportion of total frequency as does the χ^2 ellipse in the (z_1, z_2) plane.

5. METHOD OF FITTING

The most straightforward way of fitting a joint distribution of the kind proposed above seems to be as follows:

(i) Fit each marginal distribution by a curve from one of the four systems (Normal, S_L, S_B, S_U).

(ii) Calculate the observed correlation between the transformed variables, using the results of (i) to obtain the latter.

The functions f_I and f_J, and the values of all the parameters except ρ, are determined by (i); (ii) determines ρ.

The example below illustrates this method of fitting. It will be noted that in this example the higher order moments of the arrays are not well reproduced by the fitted surface. It is well to remember, however, that Pretorius (1930) obtained results which were scarcely as good when he fitted a surface of Jørgensen's Type AA to the same data. The equation of this surface contains fourteen arbitrary constants, while the S_{UU} surface fitted in the example below depends on only nine arbitrary constants.

Example. In examples 3 and 4 of Johnson (1949), curves of system S_U have been fitted to distributions of length and breadth, respectively, of 9440 beans. We now proceed to fit a surface of system S_{UU} to the joint distribution of length and breadth. The correlation between the transformed variables

$$\sinh^{-1}\left(\frac{\text{length} - 16{\cdot}0745}{1{\cdot}5192}\right) \quad \text{and} \quad \sinh^{-1}\left(\frac{\text{breadth} - 8{\cdot}6195}{0{\cdot}9721}\right)$$

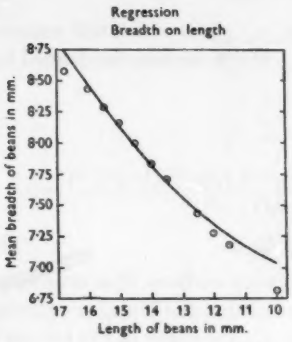

Fig. 10.

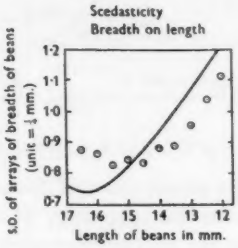

Fig. 11.

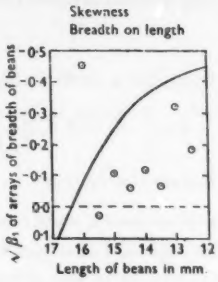

Fig. 12.

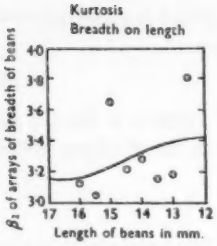

Fig. 13.

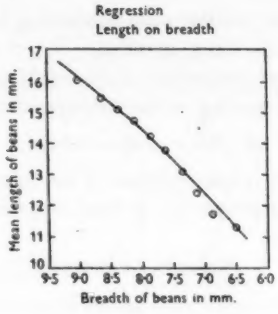

Fig. 14.

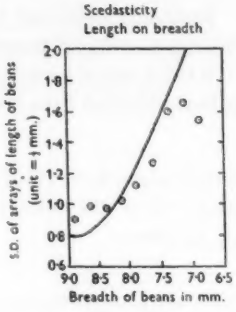

Fig. 15.

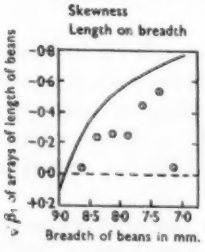

Fig. 16.

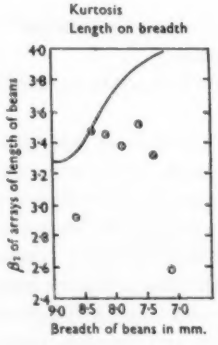

Fig. 17.

is found to be 0·746. Hence, denoting length by x_1 and breadth by x_2, the constants of the fitted S_{UV} surface are

$$\gamma_1 = 2{\cdot}38, \quad \delta_1 = 2{\cdot}64, \quad \xi_1 = 16{\cdot}0745, \quad \lambda_1 = 1{\cdot}5192,$$
$$\gamma_2 = 2{\cdot}13, \quad \delta_2 = 3{\cdot}55, \quad \xi_2 = 8{\cdot}6195, \quad \lambda_2 = 0{\cdot}9721,$$
$$\rho = 0{\cdot}746.$$

Consider now the array distribution of x_2 (breadth) for a fixed value of x_1 (length). This will be a curve of system S_U with parameters

$$\gamma(x_2 \mid x_1) = \left\{\gamma_2 - \rho\gamma_1 - \rho\delta_1 \sinh^{-1}\left(\frac{x_1 - \xi_1}{\lambda_1}\right)\right\}(1-\rho^2)^{-\frac{1}{2}}, \quad \delta(x_2 \mid x_1) = \delta_2(1-\rho^2)^{-\frac{1}{2}}.$$

Hence, inserting numerical values,

$$\delta(x_2 \mid x_1) = 5{\cdot}3307$$

and

$$\begin{aligned}\Omega(x_2 \mid x_1) &= \gamma(x_2 \mid x_1)/\delta(x_2 \mid x_1) \\ &= \delta_2^{-1}(\gamma_2 - \rho\gamma_1) - \rho\delta_1\delta_2^{-1}\sinh^{-1}\left(\frac{x_1 - 16{\cdot}0745}{1{\cdot}5192}\right) \\ &= 0{\cdot}104 - 0{\cdot}5548\sinh^{-1}\left(\frac{x_1 - 16{\cdot}0745}{1{\cdot}5192}\right).\end{aligned}$$

Also, of course, $\xi(x_2 \mid x_1) = \xi_2 = 8{\cdot}6195, \quad \lambda(x_2 \mid x_1) = \lambda_2 = 0{\cdot}9721.$

Using these values of the parameters the expected value, standard deviation, β_1 and β_2 of the array distributions of breadth were calculated. The results are shown graphically in the continuous curves in Figs. 10–13. The ringed dots show the corresponding values calculated from the observed array distributions by Pretorius. Figs. 10–13 are to be compared with that author's diagrams VI (*a*)–(*d*); to facilitate the comparison they have been constructed in a similar manner.

For array distributions of x_1 (length) for a fixed x_2 it was found that

$$\delta(x_1 \mid x_2) = 3{\cdot}9642,$$
$$\Omega(x_1 \mid x_2) = 0{\cdot}300 - 1{\cdot}0031\sinh^{-1}\left(\frac{x_2 - 8{\cdot}6195}{0{\cdot}9721}\right),$$
$$\xi(x_1 \mid x_2) = \xi_1 = 16{\cdot}0745, \quad \lambda(x_1 \mid x_2) = \lambda_1 = 1{\cdot}5192.$$

Figs. 14–17 are analogous to Figs. 10–13 and compare curves giving the theoretical values of constants of the array distributions of x_1 with the observed values (cf. Pretorius's diagrams VI (*e*)–(*h*)).

From a study of the eight diagrams (Figs. 10–17) it appears that agreement between theory and observation becomes less satisfactory as higher moments come into consideration. The regression lines are quite close fits. The scedastic lines, while not giving a close fit, seem to be fairly reasonable, and are about as good fits as the Type AA lines shown by Pretorius. In the case of the lines showing skewness and kurtosis the most that can be said is that they do indicate the general trend and order of the observed values. A peculiar feature is that it seems that in all four cases the fit of these lines would be improved if they were displaced to the right (i.e. in the direction of decreasing length or breadth).

It should be noted that Pretorius did not calculate the $\sqrt{\beta_1}$ and β_2 coefficients for the arrays of the Type AA surface.

REFERENCES

FRECHET, M. (1941). *Skand. Aktuar.* **24**, 214.
JOHNSON, N. L. (1949). *Biometrika*, **36**, 149.
NARUMI, S. (1923). *Biometrika*, **15**, 77.
PEARSON, K. (1923). *Biometrika*, **15**, 222, 231.
PRETORIUS, S. J. (1930). *Biometrika*, **22**, 109.
RHODES, E. C. (1922). *Biometrika*, **14**, 355.
STEFFENSEN, J. F. (1922). *Skand. Aktuar.* **5**, 73.
STEFFENSEN, J. F. (1941). *Skand. Aktuar.* **24**, 1.
VAN UVEN, M. J. (1947). *Ned. Akad. Wet. Proc.* **50**, 1063, 1252.
VAN UVEN, M. J. (1948). *Ned. Akad. Wet. Proc.* **51**, 41, 191.
WICKSELL, S. D. (1917). *Ark. Mat. Ast. Fys.* **12**, no. 20.
YUAN, P. T. (1933). *Ann. Math. Statist.* **4**, 30.

A TEST FOR RANDOMNESS IN A SEQUENCE OF TWO ALTERNATIVES INVOLVING A 2×2 TABLE

By P. G. MOORE, *University College, London*

1. Introduction

Recently the power function technique has been used to examine tests of randomness applied to a sequence of two alternatives. F. N. David (1947) has considered the power function of what may be termed the 'group' test, and G. I. Bateman (1948*a*) has similarly considered the 'longest run' test. In both cases the hypothesis tested, say H_0, is that there is randomness within the sequence against the admissible alternative, H_1, of dependence of the kind found in a simple Markoff chain. We propose here to deal with a slight variation of the group test which leads to a 2×2 table. We shall then examine the power of this test against the alternative of positive dependence in the sequence of the type met with in a simple Markoff chain. This will enable us, before some set of data is collected, to find approximately how large a sample is necessary in order to be able to detect a specified degree of dependence using a given significance level.

In the sequence of alternatives we write 1 for the happening of a certain event and 0 for its negation. For example, in a sequence of tyres coming off a manufacturing plant, a 1 might indicate a tyre rejected as not being up to standard. We would be interested here in seeing whether the rejected tyres were randomly placed in the sequence or whether there was an undue clustering effect of the rejected tyres. Taking successive couplets along the sequence, the usual procedure is to represent the result in a 2×2 table using the notation given in Table 1. The value of N will obviously be one less than the number of items in the sequence and also r and m will not differ by more than unity. Our test will be concerned with the form of this table under (i) randomness and (ii) dependence of the simple Markoff type.

Table 1

		Present trial: 1	Present trial: 0	Total
Preceding trial	1	a	c	m
	0	b	d	n
Total		r	s	N

2. Presentation of data

The preceding manner of arranging the data has been used, among others, by Cochran (1938). He then compared the two proportions a/m and b/n, using the standard test for this kind of data. However, for a given sequence, the values in the 2×2 table formed are not uniquely fixed by the numbers of 1's and 0's in the sequence. To clarify the position we may consider

now the various cases arise. Suppose that in a sequence of R units there are r_1 1's and r_2 0's. Obviously $r_1+r_2=R$. Four cases may now be distinguished.

A. Sequence begins with a 1 and ends with a 0.
B. Sequence begins with a 1 and ends with a 1.
C. Sequence begins with a 0 and ends with a 0.
D. Sequence begins with a 0 and ends with a 1.

A and D must each have an even number of groups, say $2t$, in the sequence, while B and C will have an odd number, say $2t+1$. These cases give rise to the 2×2 tables shown in Table 2.

Table 2

Case A
Present trial

Preceding trial	1	0	Total
1	r_1-t	t	r_1
0	$t-1$	r_2-t	r_2-1
Total	r_1-1	r_2	$R-1$

Case B
Present trial

Preceding trial	1	0	Total
1	r_1-t-1	t	r_1-1
0	t	r_2-t	r_2
Total	r_1-1	r_2	$R-1$

Case C
Present trial

Preceding trial	1	0	Total
1	r_1-t	t	r_1
0	t	r_2-t-1	r_2-1
Total	r_1	r_2-1	$R-1$

Case D
Present trial

Preceding trial	1	0	Total
1	r_1-t	$t-1$	r_1-1
0	t	r_2-t	r_2
Total	r_1	r_2-1	$R-1$

It is clear that, although the data have been exhibited in the form of a 2×2 table, the underlying criterion is still that of the number of groups in a sequence of r_1 1's and r_2 0's. Hence we are concerned with finding the power function of a test based on the number of groups in the sequence.

3. Calculation of power function

The principles of a Markoff chain have been discussed in detail, for example, by Fréchet (1938), but the simple notation used here will be that of David (1947). If E_s represents the sth event, which may either be a 1 or a 0, then under H_0, the hypothesis of independence,

$$P\{E_s=1\}=p, \quad P\{E_s=0\}=q, \quad \text{where} \quad p+q=1.$$

Under the alternate hypothesis H_1, we have

$$\begin{aligned} P\{E_1=1\} &= P, & P\{E_1=0\} &= Q, \\ P\{E_s=1 \mid E_{s-1}=1\} &= p_1, & P\{E_s=0 \mid E_{s-1}=1\} &= q_1, \\ P\{E_s=1 \mid E_{s-1}=0\} &= p_2, & P\{E_s=0 \mid E_{s-1}=0\} &= q_2, \end{aligned}$$

where $$P+Q=1, \quad p_1+q_1=1 \quad \text{and} \quad p_2+q_2=1.$$

Making the simplifying assumption (David, 1947, p. 337) that if nothing is known about the $s-1$ trials preceding the sth trial, then $P\{E_s=1\}=P$ and $P\{E_s=0\}=Q$, we have that $p_2=Pq_1/Q$ or $P=p_2/(p_2+q_1)$. Using these relations and taking, say, the case of an even number of groups, it is possible with the aid of the formulae for $P\{2t \text{ groups} \mid r_1r_2H_1\}$ given by David (1947, p. 337) to evaluate a conditional power function. However, laborious calculations are necessary to evaluate this, even for a small sequence, so that some form of approximation is required. We will find, separately for an even $(2t)$ or odd $(2t+1)$ number of groups, the forms of $p(t \mid r_1r_2H_0)$ and $p(t \mid r_1r_2H_1)$, where $p(\ )$ indicates a continuous probability distribution taken as an approximation to the set of discrete probabilities denoted by $P\{\ \}$. Then our 'conditional' critical region ω_α will be such that

$$P\{E \epsilon \omega_\alpha \mid r_1r_2H_0\} \leqslant \alpha,$$

where E is the sample point. Since our alternate hypothesis is that there is positive dependence in the sequence, we are interested in establishing the significance of low values of t. Hence we want to find the greatest value of t_α such that

$$P\{t \leqslant t_\alpha \mid H_0\} \leqslant \alpha.$$

This is equivalent to finding the greatest value of t_α satisfying

$$\sum_{t=1}^{t=t_\alpha} P\{t \mid r_1r_2H_0\} \leqslant \alpha. \tag{1}$$

The power of the test, π, will be given by

$$\begin{aligned} \pi &= P\{E \epsilon \omega_\alpha \mid r_1r_2H_1\} \\ &= P\{t \leqslant t_\alpha \mid H_1\} \\ &= \sum_{t=1}^{t=t_\alpha} P\{t \mid r_1r_2H_1\}. \end{aligned} \tag{2}$$

4. Approximation to the distribution

(a) *Even number of groups*

We will consider in the first place the case of an even number of groups. It is known (David, 1947, p. 335) that if there are an even number of groups $(2t)$, then

$$\left.\begin{aligned} P\{t \mid r_1r_2H_0\} &= k\,{}^{r_1-1}C_{t-1}\,{}^{r_2-1}C_{t-1}, \\ P\{t \mid r_1r_2H_1\} &= k'\,{}^{r_1-1}C_{t-1}\,{}^{r_2-1}C_{t-1}\left(\frac{p_2q_1}{q_2p_1}\right)^t, \end{aligned}\right\} \tag{3}$$

k and k' being constants such that $\Sigma P\{2t\}=1$, where Σ denotes summation over all possible values of t. From Wald & Wolfowitz (1940) we have that as r_1 tends to infinity, r_1/r_2 remaining finite, the distribution in (3) under H_0 tends to normality so that

$$p(t \mid H_0) = \frac{1}{\sqrt{(2\pi)\sigma}} \exp\left[-\frac{1}{2}\left(\frac{t-\tau}{\sigma}\right)^2\right], \tag{4}$$

where τ and σ^2 are the mean and variance of t, and can be derived from (3). Under H_1 we shall have

$$p(t \mid H_1) = k_1 \frac{1}{\sqrt{(2\pi)\sigma}} \exp\left[-\frac{1}{2}\left(\frac{t-\tau}{\sigma}\right)^2\right] \exp\left[t \log_e \frac{p_2q_1}{q_2p_1}\right];$$

on completing the square for t and making the integral over all possible values of t equal to unity, this gives

$$p(t \mid H_1) = \frac{1}{\sqrt{(2\pi)}\sigma} \exp\left[-\frac{1}{2\sigma^2}\left\{t - \left(\tau + \sigma^2 \log_e \frac{p_2 q_1}{p_1 q_2}\right)\right\}^2\right]. \tag{5}$$

Hence, t is again normally distributed with the same variance as before but with a new mean, τ', where

$$\tau' = \tau + \sigma^2 \log_e \frac{p_2 q_1}{p_1 q_2}. \tag{6}$$

From equations (1) and (4) we can obtain the critical region and from (2) and (5) the power. The values of τ and σ^2 may be obtained in a simple way following the method outlined for the complete distribution by Wald & Wolfowitz. We need the following three quantities:

$$\alpha = \sum_{c=1}^{r-1} {}^{r_1-1}C_{c-1}\, {}^{r_2-1}C_{c-1}, \quad \beta = \sum_{c=1}^{r_1-1} c\, {}^{r_1-1}C_{c-1}\, {}^{r_2-1}C_{c-1}, \quad \gamma = \sum_{c=1}^{r_1-1} c^2\, {}^{r_1-1}C_{c-1}\, {}^{r_2-1}C_{c-1}.$$

It will be seen that α is the term independent of x in the expansion of $(1+x)^{r_1-1}(1+1/x)^{r_2-1}$, and hence is equal to ${}^{r_1+r_2-2}C_{r_2-1}$, while $(\beta-\alpha)$ is the term independent of x in the expansion of $(r_1-1)(1+x)^{r_1-2}x(1+1/x)^{r_2-1}$, and hence is equal to $(r_1-1)\,{}^{r_1+r_2-3}C_{r_1-2}$. The expression $\sum_{c=1}^{r_1-1}(c-1)^2\, {}^{r_1-1}C_{c-1}\, {}^{r_2-1}C_{c-1}$ is the term independent of x in the expansion of

$$(r_1-1)(r_2-1)(1+x)^{r_1-2}(1+1/x)^{r_2-2}$$

and is thus equal to $(r_1-1)(r_2-1)\,{}^{r_1+r_2-4}C_{r_2-2}$. Using the identity $c^2 = (c-1)^2 + 2c - 1$ we can evaluate γ. From these values we obtain for τ and σ^2

$$\tau = \frac{r_1 r_2 - 1}{r_1 + r_2 - 2}, \tag{7}$$

$$\sigma^2 = \frac{(r_1-1)^2(r_2-1)^2}{(r_1+r_2-2)^2(r_1+r_2-3)}. \tag{8}$$

(*b*) *Odd number of groups*

For the case of an odd number of groups the expressions concerned are not nearly as simple. We have, if there be $2t+1$ groups,

$$\left.\begin{aligned} P\{t \mid r_1 r_2 H_0\} &= l({}^{r_1-1}C_t\, {}^{r_2-1}C_{t-1} + {}^{r_1-1}C_{t-1}\, {}^{r_2-1}C_t), \\ P\{t \mid r_1 r_2 H_1\} &= l'\left(\frac{p_2 q_1}{p_1 q_2}\right)^t \left[\frac{P}{p_1}\, {}^{r_1-1}C_t\, {}^{r_2-1}C_{t-1} + \frac{Q}{q_2}\, {}^{r_1-1}C_{t-1}\, {}^{r_2-1}C_t\right], \end{aligned}\right\} \tag{9}$$

l and l' being constants, such that the summation over all possible values of t is unity. We can deduce the mean (λ) and variance (μ^2) of t under H_0 from (9) by similar methods to those employed above, giving

$$\lambda = \frac{(r_1-1)^2 + (r_2-1)^2}{r_1^2 + r_2^2 - r_1 - r_2} \frac{r_1 r_2}{r_1 + r_2 - 2}, \tag{10}$$

$$\mu^2 = r_1 r_2 \frac{(r_1-1)^2(r_1 r_2 - r_2 - 1) + (r_2-1)^2(r_1 r_2 - r_1 - 1)}{(r_1+r_2-2)(r_1+r_2-3)(r_1^2+r_2^2-r_1-r_2)} - \lambda^2. \tag{11}$$

Again, following Wald & Wolfowitz, we deduce that t will tend to be normally distributed as r_1 tends to infinity, r_1/r_2 remaining finite, so that

$$p(t \mid H_0) = \frac{1}{\sqrt{(2\pi)}\mu} \exp\left[-\frac{1}{2}\left(\frac{t-\lambda}{\mu}\right)^2\right].$$

Under H_1 we have that

$$p(t \mid H_1) = \frac{1}{\sqrt{(2\pi)}\mu_1} \exp\left[-\frac{1}{2\sigma^2}\left\{t-\left(\xi+\mu_1^2 \log_e \frac{p_2 q_1}{q_2 p_1}\right)\right\}^2\right], \tag{12}$$

where ξ and μ_1^2 are the mean and variance of

$$\phi = \frac{P}{p_1}{}^{r_1-1}C_i\, {}^{r_2-1}C_{i-1} + \frac{Q}{q_2}{}^{r_1-1}C_{i-1}\, {}^{r_2-1}C_i.$$

Using the same methods as before we obtain the expressions

$$\xi = \frac{r_1 r_2}{r_1+r_2-2}\left[\frac{\dfrac{P}{p_1}(r_1-1)^2+\dfrac{Q}{q_2}(r_2-1)^2}{\dfrac{P}{p_1}r_1(r_1-1)+\dfrac{Q}{q_2}r_2(r_2-1)}\right], \tag{13}$$

$$\mu_1^2 = \frac{r_1^2 r_2^2}{(r_1+r_2-2)(r_1+r_2-3)}\left[\frac{\dfrac{P}{p_1}\dfrac{(r_1-1)^2}{r_1 r_2}(r_1 r_2-r_2-1)+\dfrac{Q}{q_2}\dfrac{(r_2-1)^2}{r_1 r_2}(r_1 r_2-r_1-1)}{\dfrac{P}{p_1}r_1(r_1-1)+\dfrac{Q}{q_2}r_2(r_2-1)}\right]-\xi^2. \tag{14}$$

5. Equivalence of the tests

We desire to show that in large samples the cases of odd and even numbers of groups tend to equivalence in the sense that the distribution of t in both cases becomes identical. As we have indicated that in both cases the distribution of t tends to normality, it is only necessary to show that the means and variances agree. Table 3 shows the scheme of means and variances that we have obtained. Hence we require to show, for large values of r_1 and r_2, that (i) λ and ξ tend to τ and (ii) μ^2 and μ_1^2 tend to σ^2. We note that, in the particular case $r_1 = r_2 = r$, the expressions reduce to

$$\tau = \tfrac{1}{2}(r+1), \quad \lambda = \xi = \tfrac{1}{2}r, \quad \sigma^2 = (r-1)^2/4(2r-3), \quad \mu^2 = \mu_1^2 = r(r-2)/4(2r-3).$$

Table 3

	Mean of t		Variance of t, H_0 and H_1
	H_0	H_1	
Even no. of groups ($2t$)	τ	$\tau' = \tau + \sigma^2 \log_e \dfrac{p_2 q_1}{q_2 p_1}$	σ^2
Odd no. of groups ($2t+1$)	λ	$\lambda' = \xi + \mu_1^2 \log_e \dfrac{p_2 q_1}{q_2 p_1}$	μ^2

When $P = 0{\cdot}5$, $\xi = \lambda$ and $\mu_1^2 = \mu^2$ whatever r_1 and r_2. In general we see, from (7) and (10), that as r_1 and r_2 get large τ and λ tend to a common value. Similarly, from (8) and (11), σ^2 and μ^2 tend to a common value. Under H_1, however, the expressions are a little more

complicated. We see, from the forms of (7) and (13), that τ and ξ both tend to $r_1 r_2/(r_1+r_2-2)$. From (8) and (14), together with (13), we see that σ^2 and μ_1^2 both tend to

$$r_1^2 r_2^2/(r_1+r_2-2)^2(r_1+r_2-3).$$

It is difficult to compare numerically the approximate with the true results as there are so many variables, but we will take some specific cases.

(i) $P = 0{\cdot}5$, *whence* $\xi = \lambda$ *and* $\mu_1^2 = \mu^2$. $r_1+r_2 = 200$.

Table 4

	τ	$\xi=\lambda$	σ^2	$\mu_1^2=\mu^2$
$r_1=100$ $r_2=100$	50·5	50	12·4378	12·4365
$r_1=105$ $r_2=95$	50·1206	49·8763	12·3744	12·3719
$r_1=110$ $r_2=90$	49·7437	49·5050	12·1853	12·1915
$r_1=115$ $r_2=85$	49·1156	48·8859	11·8733	11·8854

(ii) $P \neq 0{\cdot}5$. $r_1+r_2 = 200$.

Table 5

$P=0{\cdot}6$				$P=0{\cdot}4$			
$r_1=120$, $r_2=80$				$r_1=80$, $r_2=120$			
$\tau=48{\cdot}2362$ $\lambda=48{\cdot}0188$				$\sigma^2=11{\cdot}4433$ $\mu^2=11{\cdot}4723$			
p_1	p_2	ξ	μ_1^2	p_1	p_2	ξ	μ_1^2
0·7	0·45	48·0257	11·4679	0·5	0·3333	48·0238	11·4689
0·75	0·375	48·0280	11·4701	0·6	0·2667	48·0273	11·4696
0·8	0·3	48·0299	11·4709	0·7	0·2000	48·0299	11·4708
0·85	0·225	48·0314	11·4766	0·8	0·1333	48·0319	11·4725
0·9	0·15	48·0327	11·4770	0·9	0·0667	48·0335	11·4738

These tables show that the differences involved do not appear to be very large when the sequence is of some length.

6. The total-group test

We have obtained in the preceding sections the power function for even and odd numbers of groups separately in a sequence of r_1 1's and r_2 0's. We will now combine these results to form what David (1947) has termed the conditional power function. That is, we will find the distribution of T, the total number of groups, in a sequence of r_1 1's and r_2 0's under H_1. We have shown that for an even $(2t)$ or odd $(2t+1)$ number of groups t tends to be normally distributed with mean τ' and variance σ^2. Now in the limit, i.e. as the sequence gets large, the probabilities of getting an even or odd number of groups will each tend to a half. This can be seen if we

consider the discrete set of probabilities graduated by a normal curve. Hence we can combine the two distributions to give a new mean, θ, and variance, ν^2, of the distribution of T,

$$\theta = \Sigma(2t)\, P\{\text{even } t\} + \Sigma(2t+1)\, P\{\text{odd } t\},$$

where the summation is taken over all possible values of t. In the limit we will get

$$\theta = 2\tau' + \tfrac{1}{2}.$$

Similarly, we know that
$$\sigma^2 = 2\Sigma(t^2)\, P\{\text{even } t\} - \tau'^2$$
$$= 2\Sigma(t^2)\, P\{\text{odd } t\} - \tau'^2.$$

Thus
$$\nu^2 = \Sigma(2t)^2\, P\{\text{even } t\} + \Sigma(2t+1)^2\, P\{\text{odd } t\} - (2\tau' + \tfrac{1}{2})^2$$
$$= 2(\sigma^2 + \tau'^2) + 2(\sigma^2 + \tau'^2) + 2\tau' + \tfrac{1}{2} - 4\tau'^2 - 2\tau' - \tfrac{1}{4}$$
$$= 4\sigma^2 + \tfrac{1}{4}.$$

Under H_0, however, we may calculate the exact value of the mean and variance of T following the method used in § 4 and given by Wald & Wolfowitz (1940, p. 151). As the sequence gets large T may be assumed to be normally distributed with

$$\text{Mean} = \frac{2r_1 r_2}{r_1 + r_2} + 1, \tag{15}$$

$$\text{Variance} = \frac{2r_1 r_2 (2r_1 r_2 - r_1 - r_2)}{(r_1 + r_2)^2 (r_1 + r_2 - 1)}. \tag{16}$$

7. The hypergeometric distribution

If we have groups, then from (3),

$$P\{2t \mid r_1 r_2 H_0\} = {}^{r_1-1}C_{t-1}\, {}^{r_2-1}C_{t-1} / {}^{r_1+r_2-2}C_{r_2-1},$$

an expression proportional to a term in a hypergeometric series and also equal to the chance of the partition in a 2×2 table with fixed margins (Table 6). Hence, the mean of $(r_1 - t)$ is

Table 6

$r_1 - t$	$t - 1$	$r_1 - 1$
$t - 1$	$r_2 - t$	$r_2 - 1$
$r_1 - 1$	$r_2 - 1$	$r_1 + r_2 - 2$

$(r_1 - 1)^2/(r_1 + r_2 - 2)$ or, mean $t = (r_1 r_2 - 1)/(r_1 + r_2 - 2)$, and the variance of $(r_1 - t)$ and therefore of t is $(r_1 - 1)^2 (r_2 - 1)^2/(r_1 + r_2 - 2)^2 (r_1 + r_2 - 3)$, results which agree with those in equations (7) and (8). These are analogous to the expressions given by Stevens (1939) for the number of groups of 1's and 2's in a circular sequence, when there must be an even number of groups. For the case of an odd number of groups we may take the two terms of (9) separately, the

first term, giving an expression proportional to the probability of $2t+1$ groups in case C and the second for case B. These expressions are also proportional to a term in a different hypergeometric series and equal to the chances of the partitions in 2×2 tables with fixed margins (Tables 6A, B).

Table 6A. *Case B*

r_1-t	t	r_1
$t-1$	r_2-t-1	r_2-2
r_1-1	r_2-1	r_1+r_2-2

Table 6B. *Case C*

r_1-t-1	$t-1$	r_1-2
t	r_2-t	r_2
r_1-1	r_2-1	r_1+r_2-2

The conditional power function obtained in the preceding sections is similar in many respects to that obtained by Patnaik (1948) in testing for the difference between two proportions when the data are put in the form of a 2×2 table. His conditional power function (pp. 162–3 of his paper) is the same as that derived here for t, remembering that he is using 'a' in the notation of our Table 1 as his criterion.

The usual procedure in the past to test for randomness in a sequence of two alternatives has been to apply the χ^2 test to Table 1. It is clear, however, that the underlying basis is the group test. The procedure now suggested is that we deduce T from Table 1. Then our test would be:

(i) Find r_1, r_2 and T from the 2×2 table formed.

(ii) Calculate mean and variance from (15) and (16).

(iii) From the normal curve tables find the probability of getting the observed value of T or less under H_0.

(iv) Reject the hypothesis of independence if this probability is less than some preassigned level α.

In large sequences the two tests may be shown to tend to equivalence, but how far the difference between the tests is important for small samples has not yet been fully investigated, although in a number of cases comparisons have been made. As an example, Table 7 gives the probabilities of having T_0 groups or less in a sequence of 40 using the different methods of evaluation. For the 'normal' group test, equations (15) and (16) have been utilized with a continuity correction and the normal curve tables. For χ^2, the probability in brackets is obtained with a continuity correction, the other probability being obtained without it. The last column gives the exact chances calculated from equations (3) and (9).

It can be seen that the exact and approximate distribution of T are very similar, but with χ^2 the use of a continuity correction appears to make all the probabilities too high. This discrepancy was noted in a number of other cases. It should also be noted that when T is odd the probability level obtained from χ^2 is dependent on whether we have case B or case C. The differences between the two might be important in cases where the probabilities were near the preassigned significance level.

Table 7. *Sequence of* 40. $r_1 = 25$, $r_2 = 15$

T_0	'Normal' groups	χ^2			Exact groups
		Cases A and D	Case B	Case C	
12	0·0066	0·0066 (0·0163)	—	—	0·0065
13	0·0162	—	0·0144 (0·0324)	0·0193 (0·0425)	0·0164
14	0·0362	0·0363 (0·0733)	—	—	0·0365
15	0·0728	—	0·0656 (0·1208)	0·0847 (0·1525)	0·0738
16	0·1330	0·1339 (0·2220)	—	—	0·1329
17	0·2206	—	0·2025 (0·3105)	0·2489 (0·3707)	0·2215

The table gives $P\{T \leqslant T_0\}$.

8. THE POWER FUNCTION; NUMERICAL COMPARISONS

All the first set of calculations are based on a sequence of 40 units. The value of P used throughout is 0·5, the significance of which is mentioned in the next section and, as we are concerned with the case of positive dependence, p_1 is taken greater than P.

Table 8 compares the values of the power function of the group test for different compositions of a sequence of 40 units. These values were calculated from the normal curve approximation, using a continuity correction. Allowing for differences in the value of α, owing

Table 8. *Power functions for sequences of* 40 ($r_2 = 40 - r_1$)

p_1	Value of r_1							
	20	21	22	23	24	25	26	27
0·5*	0·0274	0·0281	0·0315	0·0348	0·0424	0·0536	0·0150	0·0223
0·55	0·0978	0·0995	0·1068	0·1144	0·1306	0·1520	0·0541	0·0713
0·6	0·2565	0·2591	0·2688	0·2811	0·3047	0·3328	0·1513	0·1807
0·65	0·5051	0·5076	0·5150	0·5283	0·5510	0·5748	0·3335	0·3683
0·7	0·7660	0·7674	0·7694	0·7783	0·7913	0·8030	0·5838	0·6103
0·75	0·9346	0·9348	0·9340	0·9372	0·9409	0·9433	0·8207	0·8301
0·8	0·9920	0·9920	0·9915	0·9920	0·9923	0·9922	0·9579	0·9586
0·85	0·9998	0·9998	0·9997	0·9997	0·9997	0·9997	0·9966	0·9963
0·9	1·0	1·0	1·0	1·0	1·0	1·0	1·0	0·9999
κ	0·1254	0·1194	0·1031	0·0807	0·0672	0·0366	0·0211	0·0109

* When $p_1 = 0·5$, the value of the power function is the value of the first kind of error, usually denoted by α.

to the discontinuity, the values are very consistent, and the test appears to be just as powerful over the limited range considered, whether or not the sequence has equal numbers of 1's and 0's. κ, given at the bottom of the table, is the probability, under the hypothesis H_0, that in a sequence of 40 units with $P = 0{\cdot}5$ there will be just r_1 1's. We may obtain the overall power function by considering a sequence of fixed length, say 40. Then, assuming that r_1 is normally distributed with mean RP and variance $RPQ(1+\delta)/(1-\delta)$, where $\delta = p_1 - p_2$ and $R = r_1 + r_2$ (Uspensky, 1937, p. 301 following Markoff), we can use Table 8, extended slightly beyond $r_1 = 27$, to get this overall power function. We do this by weighting the tabulated powers by the probability of obtaining that division of r_1 and r_2 and then summing for all possible divisions of R. This would tell us that if we drew samples of 40 from a population for which $P = 0{\cdot}5$ and $\delta = \delta_0$, then, using this group test, the chance of detecting that $\delta \neq 0$ is the power corresponding to $p_1 = 0{\cdot}5 + \frac{1}{2}\delta_0$. These powers are given in Table 8 alongside those for a group test devised by Bateman.

Table 9. *Comparison of power functions* $r_1 + r_2 = 40$

Test used	Value of p_1								
	0·5	0·55	0·6	0·65	0·7	0·75	0·8	0·85	0·9
T, as described above	0·0329	0·1080	0·2664	0·5041	0·7512	0·9183	0·9854	0·9990	1·0000
Bateman's	0·0266	0·0952	0·2484	0·4867	0·7397	0·9138	0·9846	0·9989	1·0000

Bateman (1948*b*) has shown that when $P = 0{\cdot}5$ the total number of groups (T) is distributed as the binomial $(p_1 + q_1)^{R-1}$. Hence we may compare this test with the power function just computed. The slight differences are due to the fact that whereas Bateman's test has the same critical region, whatever r_1, our test has a critical region depending on r_1. The agreement is, however, very good. But when $P \neq 0{\cdot}5$ Bateman's formula no longer applies and our approximate test must be used.

The difficulty in getting from Table 8 to Table 9 is that it entails very long calculations. As an approximation to the overall power we may follow a method used by Patnaik (1948). Briefly, we replace r_1 and r_2 by their expected values of RP and $R(1-P)$. The adequacy of this approximation can be seen by comparing the first column in Table 8 with Table 9. The agreement is very good.

Using the 0·05 nominal significance level for H_0, we can find the minimum length of sequence such that the chance is 0·9 of rejecting randomness by the group test, when p_1 has values greater than 0·5. Table 10 gives these minimum lengths of sequence. They are obtained in the same way as before, successive values of r_1 and r_2 being taken until the power just becomes 0·9. The table shows that if we are interested in picking out the degree of dependence of a certain magnitude the length of sequence necessary to do this with a given power varies considerably according to the value of P. In practice, however, it is possible that we would want the test to be equally powerful for the same values of some function of δ and P, and not just for equal values of δ alone.

Table 10

	Value of $\Delta = p_1 - P = \delta(1-P)$							
	0·05	0·10	0·15	0·20	0·25	0·30	0·35	0·40
$P = 0·3$	1800	466	210	126	92	70	46	34
$P = 0·4$	1240	320	120	76	48	38	24	—
$P = 0·5$	904	220	96	56	34	22	—	—
$P = 0·6$	540	120	56	38	20	—	—	—
$P = 0·7$	124	63	34	23	—	—	—	—

9. APPLICATION OF RESULTS

In a control chart an examination of runs of values above and below the median is often of assistance, in eliminating trend effects. Mosteller (1941) uses the length of the runs, while Shewhart (cited by Swed & Eisenhart, 1943) suggests the use of the technique we have discussed employing the number of runs. In cases of short sequences the power of the group test is low. But for longer sequences, say 40 or over, the technique provides a useful method for picking out clustering effects. It is assumed, from previous results, that the median value is known, whence it could be used to make $P = 0·5$. The foregoing test has been concerned solely with the case of positive dependence, i.e. $p_1 \geqslant P$, and if it were desired to test for dependence in both directions a two-ended test would have to be made on similar lines.

Table 11

		Present month: Wet	Present month: Dry	Total
Preceding month	Wet	542	582	1124
	Dry	582	759	1341
	Total	1124	1341	2465

Cochran (1938), in considering the distribution of wet and dry months, gives a 2×2 table. A wet month is defined as a month when rainfall is greater than the average for that month over the period 1881–1915. Hence we have a P of approximately 0·5. The alternative to randomness in the order in which wet and dry months follow each other is that there is a persistence of one kind of weather giving us the basic conditions for a Markoff chain. Since we are considering either case B or C of § 2, the value of T is $2 \times 582 + 1 = 1165$. Further, assuming that we have case C,* so that $r_1 = 1124$ and $r_2 = 1342$, we get that mean $T = 1224·36$ and S.E. of $T = 24·63$. Hence we refer the ratio $(1224·36 - 1124)/24·63 = 2·41$ to the normal curve tables and find that the probability of getting a value of T less than or equal to 1165,

* To differentiate between cases B and C we would need to know whether the sequence started with a wet or a dry month.

under the hypothesis of independence, is 0·0080 or odds of 125 to 1 against independence. If a continuity correction is used the odds are 119 to 1 against independence. Thus we would reject the hypothesis H_0 in favour of H_1, the hypothesis that there is positive dependence, in the sense that wet months tend to follow wet months and similarly for dry months. Cochran found the proportion of months, following a wet month, which were wet and also the proportion following a dry month which were wet. These proportions are 0·4822 and 0·4340 respectively. The S.E. of the difference is, from the binomial theorem,

$$\sqrt{\{\tfrac{1341}{2465}\cdot\tfrac{1124}{2465}\,(\tfrac{1}{1124}+\tfrac{1}{1341})\}} = \pm 0{\cdot}02013,$$

and from normal curve tables we get odds of 120 to 1 against the observed difference arising by chance in a sample from a population for which these two proportions are equal. This test is equivalent to calculating $\sqrt{\chi^2}$, and if a continuity correction is employed the odds become 108 to 1 against independence. Thus using either test, in this case of a long sequence, we would reject the hypothesis of independence, although the difference due to the use of a continuity correction for χ^2, which was noticed for short sequences, is still apparent.

10. Summary

We have obtained an approximate formula which makes it easy to find the power of the group test for a sequence of two alternatives when H_0 is the hypothesis that there is randomness in the sequence and whilst for H_1 the dependence follows a simple Markoff chain. This formula has been used to find the minimum lengths of sequence necessary to pick out various degrees of dependence with a given power. Finally, the χ^2 test for independence in a 2×2 table is shown to lead to the same test for long sequences.

The author wishes to thank Dr F. N. David and Professor E. S. Pearson for their guidance during the preparation of this paper.

REFERENCES

Bateman, G. I. (1948*a*). *Biometrika*, **35**, 97.
Bateman, G. I. (1948*b*). Unpublished thesis.
Cochran, W. G. (1938). *Quart. J.R. Met. Soc.* **64**, 631.
David, F. N. (1947). *Biometrika*, **34**, 335.
Fréchet, M. (1938). *Recherches théoriques modernes sur le calcul des probabilités*, Book 2.
Mosteller, F. (1941). *Ann. Math. Statist.* **12**, 228.
Patnaik, P. B. (1948). *Biometrika*, **35**, 157.
Stevens, W. L. (1939). *Ann. Eugen., Lond.*, **9**, 10.
Swed, F. S. & Eisenhart, C. (1943). *Ann. Math. Statist.* **14**, 66.
Uspensky, J. V. (1937). *Introduction to Mathematical Probability.* McGraw Hill Book Co.
Wald, A. & Wolfowitz, J. (1940). *Ann. Math. Statist.* **11**, 147.

A GENERAL DISTRIBUTION THEORY FOR A CLASS OF LIKELIHOOD CRITERIA

By G. E. P. BOX

Imperial Chemical Industries, Dyestuffs Division Headquarters, Blackley, Manchester

1. Introduction

The likelihood ratio method of Neyman & Pearson (1928) has been used by many different workers for the derivation of criteria appropriate for the testing of a large variety of hypotheses. Plackett (1946), in a recent survey of literature on testing the equality of variances and covariances, lists, on this problem alone, criteria for the testing of no less than thirty-one hypotheses investigated at different times by workers in this field. Most of the criteria either have been or can be arrived at by the likelihood ratio method. In the preface to his survey Plackett says: 'Generally speaking the difficulties in testing such hypotheses lie not so much in deriving criteria—but in finding their exact distributions when the hypotheses are true and determining the best critical region to adopt.'

Although in many cases the exact distributions cannot be obtained in a form which is of practical use, it is usually possible to obtain the moments, and these may be used to obtain approximations. In some cases, for instance, a suitable power of the likelihood statistic has been found to be distributed approximately in the type I form, and good approximations have been obtained by equating the moments of the likelihood statistic to this curve. For example, in the original paper on the L_1 test for homogeneity of variances, Neyman & Pearson (1931) suggested that the distribution could be approximately represented in this way, and later Bishop & Nair (1939) showed that the significance points obtained by Nayer (1936), using this method, were in excellent agreement with the true values. The fitting of the type I curve is simple once the moments are obtained, but these moments, being the products of Γ-functions, are usually rather troublesome to calculate. To overcome this difficulty, Bishop (1939), working on the distribution of the multivariate equivalent of the L_1 test (the test for constancy of variances and covariances in k p-variate samples), derived empirical expressions for the parameters of the appropriate type I curve, thus avoiding the troublesome intermediate step of calculating moments. Bishop mentions that Nair succeeded in finding similar expressions on a theoretical basis, and Tukey & Wilks (1946) give a more general theoretical method to find approximations of this kind.

A different line of approach was adopted by Bartlett (1937). Neyman & Pearson had pointed out in their original paper that, if N' is the total sample size, $-N' \log_e L_1$ would be asymptotically distributed as χ^2. From considerations of sufficiency Bartlett obtained what was in effect a modified form (which, following Hartley & Pearson (1946), we shall call M) of this logarithmic statistic. From the moments of the modified likelihood statistic he was able to develop a scale factor C, which was related to the effective sizes of samples and which approached the value unity as the sample sizes became large. The distribution of M/C was then very well represented by χ^2 even when the samples were small. Bartlett later (1938) used the same method to obtain an approximation for the test of significance in multivariate analysis. In 1940 Hartley, starting from the moments of the modified likelihood statistic,

obtained an asymptotic series of χ^2 integrals for the logarithmic statistic M which agreed very closely with the exact distribution. In 1941 Wald & Brookner, investigating an entirely different problem, the distribution of Wilks's statistic for testing independence of k sets of variates, again starting from the moments of the likelihood statistic λ, eventually obtained an expression for the distribution of a logarithmic statistic (a negative multiple of $\log_e \lambda$) in the form of an asymptotic χ^2 series. This was later modified by Rao (1948) in the important special case of two groups, when it corresponds to the test of significance in multivariate analysis previously referred to. Neither Wald & Brookner nor Rao investigated the accuracy of these series.

It is possible therefore to distinguish two definite lines of approach, which have been used in certain cases where the moments of the likelihood criteria are known but the exact distributions are not. On the one hand the moments have been used to fit the Pearson-type curve. This usually gives an adequate approximation, but owing to the amount of labour involved in the calculation of the moments it would not be attractive for routine significance testing unless methods such as Bishop's could be used to obtain the parameters of the fitted curve directly, or the results from the method could be tabled. On the other hand, the general expression for the moments of the likelihood statistic has been used in certain cases to obtain for the distribution of the logarithmic statistic M, a χ^2 approximation and an asymptotic χ^2 series. It will be the object of the present paper to investigate in some detail this second line of attack.

The method will be investigated in particular for two general criteria:

(1) The test of constancy of variance and covariance of k sets of p-variate samples. This includes, as an important special case when $p = 1$, the test for constancy of variance in k samples.

(2) Wilks's test for the independence of k sets of residuals, the lth set having p_l variates. When $k = 2$ this corresponds to the test of significance used in multivariate regression and analysis of variance and covariance, and when $k = 2$ and p_1 or p_2 is unity, it gives the corresponding well-known univariate tests. In the latter case, of course, the exact distributions are known.

We shall refer to these two criteria as generalized tests for homoscedasticity and independence, respectively. The assumption of normality or multinormality for the distributions of the original observations will be made throughout this paper.

Taking for our test function M, a negative multiple of the natural logarithm of the likelihood statistic (or some modification of it), we shall obtain in each case,

(a) a series solution which, we shall demonstrate, agrees very closely with the exact distribution,

(b) an approximate solution using a single χ^2 distribution,

(c) a rather better approximation using a single F distribution.

The accuracy of the various methods and the relation of the results to those of other workers will be discussed.

2. The generalized test of homoscedasticity

The univariate statistic. The L_1 statistic of Neyman & Pearson for testing the homogeneity of a set of variances, takes the form of the ratio of a weighted geometric mean of variances to a weighted arithmetic mean, where the weights are the sample numbers. Welch (1935,

1936) generalized the test to cover the case when residuals from a fitted regression equation were tested for homoscedasticity, and derived the moments for a modified criterion in which the weights could have any values whatever. In 1936 Nayer tabled the approximate significance points for L_1 in the cases of equal sample numbers by fitting type I curves to the distributions by the method of moments, as suggested in the original memoir by Neyman & Pearson.

The statistic proposed by Bartlett (1937) which we shall call M is given by

$$M = N \log_e s - \sum_l \nu_l \log_e s_l,$$

where

$$s = (\Sigma \nu_l s_l)/N,$$

and s_l is the usual unbiased estimate of the variance in the lth group, $l = 1, 2, \ldots, k$, based on sums of squares having ν_l degrees of freedom, and $N = \Sigma \nu_l$. It was later shown (Brown, 1939; Pitman, 1939; Bishop & Nair, 1939) that this criterion is unbiased in the sense used by Neyman & Pearson (1936, 1938). Nair (1939) derived a series solution for the distribution of the likelihood statistic in the case of equal sample numbers; his solution is very involved, but has been used as a standard to check approximations. Bishop & Nair (1939) used this series to check the accuracy of the type I approximation used in Nayer's table. They also checked the Bartlett (1937) approximation and found that both methods were fairly good except when the degrees of freedom were small. In the case of unequal samples, however, Nayer's tables were not available, and in view of the labour involved in the type I fit, the χ^2 method of Bartlett's was preferred.

Hartley's (1940) asymptotic series depended to the degree of approximation used, on two parameters c_1 and c_3 which varied with the effective sample size and relative composition of the groups

$$c_1 = \Sigma \frac{1}{\nu_l} - \frac{1}{N}, \quad c_3 = \Sigma \frac{1}{\nu_l^3} - \frac{1}{N^3}.$$

The first is related to Bartlett's scale factor C, in fact

$$C = 1 + \frac{c_1}{3(k-1)}.$$

Tables were afterwards computed by Thompson & Merrington (1946) from Hartley's formula, and comparisons were made with the values calculated by Bishop & Nair.

The multivariate statistic. In the multivariate case Wilks (1932) derived the likelihood ratio test and obtained the moments of the criterion, which is a generalized form of that used in the univariate test, the determinants of estimated variances and covariances replacing the variances. Bishop (1939) took as his criterion l_1, the $1/N'$th power of the likelihood statistic, N' being the total number of observations. He gave reasons for believing that this criterion could be approximately represented by a type I curve

$$p(l_1) = \text{constant}\, l_1^{m_1 - 1}(1 - l_1)^{m_2 - 1}, \tag{1}$$

by choosing the value of m_1 and m_2 so that the first two moments of the Pearson curve agreed with those of the criterion. His arguments were supported by the agreement found in a number of trials between the higher moments of the fitted type I curve and those of the criterion. Only in the case of two groups and either one or two variates was it possible to obtain a check against the exact distribution, but in these cases the agreement was very good. Unfortunately the labour involved in the calculation of the first two moments of the criterion was too

great to allow this method to be recommended for routine use. Bishop therefore proceeded as follows:

(*a*) For the case of equal sample sizes he obtained, empirically, expressions for m_1 and m_2 in terms of the number of observations n in each group, the number of variates p, and the number of groups k

$$\left.\begin{aligned} m_1 &= k(n-p) - 0{\cdot}01(k-1)(90-39p+9p^2), \\ m_2 &= 0{\cdot}25(k-1)p(p+1). \end{aligned}\right\} \tag{2}$$

(*b*) For unequal sample sizes he proposed approximating to $-2N'\log_e l_1$ by means of a χ^2 distribution using a scale factor G in a similar way to that adopted by Bartlett in the univariate case.

He showed that $-2N'\log_e l_1$ is approximately distributed as $G\chi^2$, where

$$G = 1+\frac{1}{f}\Big[\sum_{l=1}^{k}\sum_{i=1}^{p}\{i^2/2n_l+i/(n_l-i)+n_l/3(n_l-i)^2\} - \sum_{i=1}^{p}\{(k+i-1)^2/2N'+(k+i-1)/(N'-k+1-i)+N'/3(N'-k+1-i)^2\}\Big]. \tag{3}$$

n_l is the number of observations in the lth group, $\sum_l n_l = N'$ and χ^2 is distributed with $f = \frac{1}{2}(k-1)p(p+1)$ degrees of freedom.

We shall refer to these methods as Bishop's methods (*a*) and (*b*). Bishop remarks that the scale factor G is rather troublesome to calculate unless $n_l = n$. George (1945) was able to evaluate the exact distribution in a number of simple cases. She used her results to test the accuracy of Bishop's approximations and found, in the cases she considered, that (*b*) was superior to (*a*).

Plackett (1947) suggested that in view of the unsatisfactory position with regard to the distribution of this criterion that it might be better to abandon it in favour of an alternative test derived by him which had the advantage that at least when $p = 1$ or 2, and for certain other special cases, the exact distribution was known. Plackett's test, however, has the disadvantages that the results depend on the particular arrangement of observations chosen, and that the samples must be of equal sizes.

2·1. *The present approach*

Suppose s_{ijl} is the usual unbiased estimate of the variance or covariance A^{ijl} between the ith and jth variable in the lth sample based on sums of squares and products having ν_l degrees of freedom, and suppose there are k such samples and s_{ij} is the average variance or covariance $\left(\sum_l \nu_l s_{ijl}\right)\Big/N$, where $N = \sum_l \nu_l$. We take as our criterion a generalized form of Bartlett's criterion

$$M = N\log_e|s_{ij}| - \sum_l(\nu_l\log_e|s_{ijl}|) \tag{4}$$

$$= -N\log_e L_1', \quad \text{where} \quad L_1' = \prod_{l=1}^{k}\left\{\frac{|s_{ijl}|}{|s_{ij}|}\right\}^{\nu_l/N} \tag{5}$$

When the degrees of freedom are equal, M and L_1' are related with Bishop's l_1 as follows:

$$M = -2N\log_e l_1 \quad \text{and} \quad L_1' = l_1^2. \tag{6}$$

When the degrees of freedom are unequal, L_1' will differ from the likelihood statistic in weighting. When $p = 1$, M is the criterion derived by Bartlett and later used by Hartley.

We proceed to obtain the moments of L_1' when the null hypothesis is true. If c_{ijl} are the sums of squares and products based on ν_l degrees of freedom corresponding to the s_{ijl}, we have $c_{ijl} = \nu_l s_{ijl}$, $c_{ij} = N s_{ij}$, so that $c_{ij} = \sum_l c_{ijl}$. The joint probability density of the c_{ijl} for the lth sample is given by the distribution discovered by Wishart (1928):

$$p(c_{11l}, c_{12l}, \ldots, c_{ppl}) = K(\nu_l)\,|\,c_{ijl}\,|^{\frac{1}{2}(\nu_l - p - 1)} \exp\{-\tfrac{1}{2}\sum_{ij} A_{ijl} c_{ijl}\}, \tag{7}$$

where

$$\{K(\nu_l)\}^{-1} = 2^{\frac{1}{2}\nu_l p}\,\pi^{\frac{1}{4}p(p-1)} \prod_{j=0}^{p-1} \Gamma\left(\frac{\nu_l - j}{2}\right) |\,A_l\,|^{-\frac{1}{2}\nu_l}, \tag{8}$$

and A_l is the matrix of the A_{ijl}, the inverse of the matrix of the A^{ijl}.

When the null hypothesis is true, A_l is the same for each of the samples, $A_l = A$, $l = 1, 2, \ldots, k$, and the gth moment of $|\,c_{ij}\,|$ is

$$\int \prod_{l=1}^{k} \left\{K(\nu_l)\,|\,c_{ijl}\,|^{\frac{1}{2}(\nu_l - p - 1)}\,|\,c_{ij}\,|^{g/k} \exp\left(-\tfrac{1}{2}\sum_{ij} A_{ij} c_{ijl}\right)\right\} dc_{111} dc_{121} \ldots dc_{ppk}; \tag{9}$$

it is also given by

$$\int K(N)\,|\,c_{ij}\,|^{\frac{1}{2}(N - p - 1)}\,|\,c_{ij}\,|^{g} \exp\left\{-\tfrac{1}{2}\sum_{ij} A_{ij} c_{ij}\right\} dc_{11} dc_{22} \ldots dc_{pp}. \tag{10}$$

Writing $\nu_l(1+2h)$ for ν_l on both sides of the identity and then taking $g = -Nh$ and integrating over the whole space for which the matrices of the c_{ij}, c_{ijl} are positive definite, we have

$$\mathscr{E}\left[\prod_{l=1}^{k} \left(\frac{|\,c_{ijl}\,|}{|\,c_{ij}\,|}\right)^{\nu_l}\right]^h \prod_{l=1}^{k}\left[\frac{K\{\nu_l(1+2h)\}}{K(\nu_l)}\right] = \frac{K\{N(1+2h)\}}{K(N)}. \tag{11}$$

That is

$$\mathscr{E}(L_1')^{Nh} = \mathscr{E}(e^{-M})^h = \frac{K\{N(1+2h)\}}{K(N)} \prod_{l=1}^{k}\left[\frac{K(\nu_l)}{K\{\nu_l(1+2h)\}}\left(\frac{N}{\nu_l}\right)^{p\nu_l h}\right] \tag{12}$$

$$= \prod_{l=1}^{k}\left(\frac{N}{\nu_l}\right)^{h\nu_l p} \prod_{j=0}^{p-1}\left[\frac{\Gamma\{\frac{1}{2}(N-j)\}}{\Gamma\{\frac{1}{2}(N(1+2h)-j)\}} \prod_{l=1}^{k} \frac{\Gamma\{\frac{1}{2}(\nu_l(1+2h)-j)\}}{\Gamma\{\frac{1}{2}(\nu_l - j)\}}\right]. \tag{13}$$

We have first proved equation (13) as an analytic identity for real h; it will, however, be generally valid in the range where the functions are analytic. We can thus obtain an expression for the characteristic function of ρM, where ρ is a constant $\leqslant 1$ at our choice, by replacing h by $-it\rho$ in the above expression. The reason for introducing the constant ρ will appear later. Further, if we write $N = \nu k$ (i.e. ν is the average of the degrees of freedom) and define new quantities μ, μ_l, β, β_l by the relations

$$\mu_l = \rho\nu_l, \quad \mu = \rho\nu, \quad \nu = \mu + \beta, \quad \nu_l = \mu_l + \beta_l, \tag{14}$$

we obtain the characteristic function of ρM in the form

$$\Phi(t) = \prod_{l=1}^{k}\left(\frac{k\mu}{\mu_l}\right)^{-it p \mu_l} \prod_{j=0}^{p-1}\left[\frac{\Gamma[\frac{1}{2}\{k\mu + k\beta - j\}]}{\Gamma[\frac{1}{2}\{k\mu(1-2it) + k\beta - j\}]} \prod_{l=1}^{k} \frac{\Gamma[\frac{1}{2}\{\mu_l(1-2it) + \beta_l - j\}]}{\Gamma[\frac{1}{2}(\mu_l + \beta_l - j)]}\right], \tag{15}$$

and taking logarithms we have the cumulant generating function in the form

$$\Psi(t) = g(t) - g(0), \tag{16}$$

where

$$g(t) = -\sum_{l=1}^{k} it\mu_l p \log\left(\frac{k\mu}{\mu_l}\right) + \sum_{j=0}^{p-1}\left[\sum_{l=1}^{k}\{\log\Gamma[\tfrac{1}{2}\{\mu_l(1-2it) + \beta_l - j\}]\} - \log\Gamma[\tfrac{1}{2}\{k\mu(1-2it) + k\beta - j\}]\right], \tag{17}$$

and $g(0)$ is a constant independent of t obtained by putting $t = 0$ in the above expression. Now Barnes (1899) was able to generalize Stirling's theorem, and he showed that for all x, real or complex, $\log \Gamma(x+h)$ may be expanded in an asymptotic series:

$$\log \Gamma(x+h) = \log \sqrt{(2\pi)} + (x+h-\tfrac{1}{2}) \log x - x - \sum_{r=1}^{n} (-1)^r \frac{B_{r+1}(h)}{r(r+1)\, x^r} + R_{n+1}(x), \tag{18}$$

where $R_m(x)$ is a remainder term such that $|R_m(x)| \leqslant \dfrac{\theta}{|x^m|}$, θ is some constant independent of x and $B_r(h)$ is the Bernoulli polynomial of degree r and order unity defined by

$$\frac{\tau e^{h\tau}}{e^{\tau} - 1} = \sum_{r=0}^{\infty} \frac{\tau^r}{r!} B_r(h). \tag{19}$$

Expanding each of the Γ-functions in this manner we obtain

$$\Psi(t) = Q - g(0) - \tfrac{1}{4}(k-1)\, p(p+1) \log(1-2it) + \sum_{r=1}^{n} \frac{\alpha_r}{\mu^r} (1-2it)^{-r} + R_{n+1}(\mu, t), \tag{20}$$

where Q does not contain t and is given by

$$Q = \frac{p(k-1)}{2} \log 2\pi + \frac{p}{2}\left[\sum_{l=1}^{k} \left(\nu_l - \frac{p+1}{2} \right) \log \frac{\mu_l}{2} - \left(k\nu - \frac{p+1}{2} \right) \log \frac{k\mu}{2} \right], \tag{21}$$

$$\alpha_r = \frac{(-1)^{r+1} (2\mu)^r}{r(r+1)} \sum_{j=0}^{p-1} \left[\sum_{l=1}^{k} \frac{B_{r+1}\left(\frac{\beta_l - j}{2}\right)}{\mu_l^r} - \frac{B_{r+1}\left(\frac{k\beta - j}{2}\right)}{\mu^r k^r} \right], \tag{22}$$

and $R_{n+1}(\mu, t)$ is defined by (17) and (18).

From (20) we have

$$\Phi(t) = K(1-2it)^{-\frac{1}{2}f} \sum_{v=0}^{\infty} \frac{a_v}{\mu^v} (1-2it)^{-v} + R'_{n+1}(\mu, t) \tag{23}$$

where $K = \exp\{Q - g(0)\}$, $f = \frac{1}{2}(k-1)\, p(p+1)$, and a_v is the coefficient of μ^{-v} in the expansion of $\exp\left\{ \sum_{r=1}^{n} \alpha_r / \mu^r \right\}$.

The probability density function of ρM is then given by

$$p(\rho M) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-it\rho M} \Phi(t)\, dt \tag{24}$$

$$= K \sum_{v=0}^{\infty} \frac{a_v}{\mu^v} p(\chi^2_{f+2v}) + R''_{n+1}(\mu, t). \tag{25}$$

The probability that a given value M_0 of the criterion is exceeded is therefore

$$\Pr.\{M \geqslant M_0\} = K \sum_{v=0}^{\infty} \frac{a_v}{\mu^v} P_{f+2v} + R'''_{n+1}(\mu, t), \tag{26}$$

where

$$P_{f+2v} = \int_{\rho M_0}^{\infty} p(\chi^2_{f+2v})\, d\chi^2, \tag{27}$$

$$R'''_{n+1}(\mu, t) = \frac{K}{2\pi} \int_{\rho M_0}^{\infty} \int_{-\infty}^{\infty} e^{-it\rho M} \sum_{v=0}^{\infty} \frac{a_v}{\mu^v} (1-2it)^{-\frac{1}{2}(f+2v)} \left(e^{R_{n+1}(\mu, t)} - 1\right) dt\, d(\rho M)$$

and $p(\chi^2_{f+2v})$ is the probability density function of the χ^2 distribution with $f+2v$ degrees of freedom. For all sufficiently large values of μ, $R'''_{n+1}(\mu, t)$ tends to zero and the required

probability will be given with sufficient accuracy by taking a suitable number of terms of the series in (26). Putting $t = 0$ in (20) we have

$$Q - g(0) = -\left\{\sum_{r=1}^{n} (\alpha_r/\mu^r) + R_{n+1}(\mu, 0)\right\}. \tag{28}$$

It is found in practice that by taking a few terms of the series (even in difficult cases usually not more than six), $\exp(-\Sigma\alpha_r/\mu_r)$ is so close to $\exp\{Q - g(0)\} = K$ that direct calculation of that constant is unnecessary.

If we expand $Q - g(0)$ as well as $g(t)$ we obtain instead of (20)

$$\Psi(t) = -\tfrac{1}{2}f\log(1-2it) + \sum_{r=1}^{n}\left\{\frac{\alpha_r}{\mu^r}(1-2it)^{-r} - 1\right\} + R_{n+1}(\mu, t) - R_{n+1}(\mu, 0). \tag{29}$$

Proceeding as before we obtain

$$\begin{aligned}\Pr.\{M \geqslant M_0\} = {} & P_f + \alpha_1(P_{f+2} - P_f)\,1/\mu \\ & + \left\{\alpha_2(P_{f+4} - P_f) + \frac{\alpha_1^2}{2!}(P_{f+4} - 2P_{f+2} + P_f)\right\} 1/\mu^2 + \text{etc.}\end{aligned} \tag{30}$$

Thus we may use a suitable number of terms of either of the series given in (26) or (30) to obtain the probability of the criterion exceeding a given value. Formula (26) has been used in this paper, (30) being rather unwieldy if a large number of terms have to be taken.

It should be noted that in the derivation we have used two series, first the asymptotic series for the expansion of the Γ-functions, and then the exponential series. In any particular case we have to decide how many terms we need in the asymptotic series to give a sufficiently close representation of the function and then how many terms we shall use in the exponential series. In those cases investigated here, six terms of the exponential series have nearly always proved adequate as judged by the closeness of agreement between $\Sigma\alpha_r/\mu^r$ and $g(0) - Q$ independently calculated; often fewer terms were necessary, terms in higher powers of $1/\mu$ having negligible effect. In the case of the exponential series the number of terms necessary to represent adequately $\exp\sum_{r}^{n}\alpha_r/\mu^r$ is usually not more than eight, but has sometimes been as many as fourteen. It is mainly in order to keep the number of terms required at this stage within manageable limits that the scale factor ρ is introduced, since by suitable choice of this constant, the values of the α's can be kept small and the number of terms required in the exponential series is consequently less. We see that in effect we are fitting a χ^2 series to the statistic M by arranging that, to the order of accuracy chosen in the asymptotic series, the series will have *all* its cumulants identical with those of M. Before we consider the problem of choosing a suitable value for ρ, we shall derive an expression for the α's and hence the a's in a form which is more suitable for computation.

2·2. *Determination of the α's in a form suitable for calculation*

From the well-known properties of Bernoulli polynomials (see, for example, Milne-Thomson, 1933), we may write the symbolic equality

$$B_r(x+y) \doteqdot (B+x+y)^r, \tag{31}$$

where, after expansion, each index of B is to be replaced by the corresponding suffix. Whence

$$B_{r+1}\left(\frac{\beta_i - j}{2}\right) = \sum_{s=0}^{r+1}\binom{r+1}{s}\left(\frac{\beta_i}{2}\right)^{r+1-s} B_s\left(\frac{-j}{2}\right). \tag{32}$$

Also if $P(x)$ is a polynomial in x and $P'(x)$ is the differential coefficient of $P(x)$ with respect to x,

$$\sum_{x=0}^{p-1} P'(x) \doteqdot P(B+p) - P(B). \tag{33}$$

Thus if $$P'(j) = B_s\left(\frac{-j}{2}\right),$$

$$P(j) = -\frac{2}{s+1} B_{s+1}\left(\frac{-j}{2}\right) + \text{constant}, \tag{34}$$

and $$\sum_{j=0}^{p-1} B_s\left(\frac{-j}{2}\right) \doteqdot \frac{-2}{s+1}\left[B_{s+1}\left\{-\frac{B+p}{2}\right\} - B_{s+1}\left\{-\frac{B}{2}\right\}\right]. \tag{35}$$

If we denote the expression in square brackets by δ_s, we obtain from (32) and (35)

$$\sum_{j=0}^{p-1}\sum_{l=1}^{k} B_{r+1}\left(\frac{\beta_l - j}{2}\right) \doteqdot -2\sum_{l=1}^{k}\sum_{s=0}^{r+1}\binom{r+1}{s}\frac{1}{s+1}\left(\frac{\beta_l}{2}\right)^{r+1-s}\delta_s, \tag{36}$$

and in a similar way we find

$$\sum_{j=0}^{p-1} B_{r+1}\left(\frac{k\beta - j}{2}\right) \doteqdot -2\sum_{s=0}^{r+1}\binom{r+1}{s}\frac{k^{r+1-s}}{s+1}\left(\frac{\beta}{2}\right)^{r+1-s}\delta_s. \tag{37}$$

Whence from (22) we obtain α_r as a polynomial of degree r in β

$$\alpha_r = \frac{(-1)^r k}{r(r+1)(r+2)}\sum_{s=1}^{r+1}\binom{r+2}{s+1} 2^s D_s \beta^{r+1-s}, \tag{38}$$

where $$D_s = \delta_s \gamma_s,$$

$$\delta_s \doteqdot B_{s+1}\left\{-\frac{B+p}{2}\right\} - B_{s+1}\left\{-\frac{B}{2}\right\}, \tag{39}$$

and $$\gamma_s = \frac{1}{k}\sum_{l=1}^{k}\left(\frac{\nu}{\nu_l}\right)^{s-1} - \frac{1}{k^s}. \tag{40}$$

It is interesting to note the relation between these quantities and the c's defined by Hartley; if

$$c_s = \sum_l \frac{1}{\nu_l^s} - \frac{1}{N^s}, \quad \gamma_s = k^{-1}\nu^{s-1}c_{s-1}. \tag{41}$$

In the special case when the samples have equal degrees of freedom

$$\gamma_s = 1 - 1/k^s.$$

The values for δ_s for $s = 0, 1, \ldots, 7$, found from equation (39), are given below:

s	δ_s
0	$-\frac{1}{2}p$,
1	$\frac{1}{6}p(p+1)$,
2	$-\frac{1}{12}p(2p^2+3p-1)$,
3	$\frac{1}{16}p(p-1)(p+1)(p+2)$,
4	$-\frac{1}{192}p(6p^4+15p^3-10p^2-30p+3)$,
5	$\frac{1}{128}p(p-1)(p+1)(p+2)(2p^2+2p-7)$,
6	$-\frac{1}{768}p(6p^6+21p^5-21p^4-105p^3+21p^2+147p-5)$,
7	$\frac{1}{768}p(p-1)(p+1)(p+2)(3p^4+6p^3-23p^2-26p+62)$,

(42)

and the values for the first six α's from equation (38) are

$$\left.\begin{aligned}
\alpha_1 &= -\tfrac{1}{3}k\{3D_1\beta + 2D_2\},\\
\alpha_2 &= \tfrac{1}{6}k\{3D_1\beta^2 + 4D_2\beta + 2D_3\},\\
\alpha_3 &= -\tfrac{1}{15}k\{5D_1\beta^3 + 10D_2\beta^2 + 10D_3\beta + 4D_4\},\\
\alpha_4 &= \tfrac{1}{60}k\{15D_1\beta^4 + 40D_2\beta^3 + 60D_3\beta^2 + 48D_4\beta + 16D_5\},\\
\alpha_5 &= -\tfrac{1}{105}k\{21D_1\beta^5 + 70D_2\beta^4 + 140D_3\beta^3 + 168D_4\beta^2 + 112D_5\beta + 32D_6\},\\
\alpha_6 &= \tfrac{1}{420}k\{7D_1\beta^6 + 28D_2\beta^5 + 70D_3\beta^4 + 112D_4\beta^3 + 112D_5\beta^2 + 64D_6\beta + 16D_7\}.
\end{aligned}\right\} \quad (43)$$

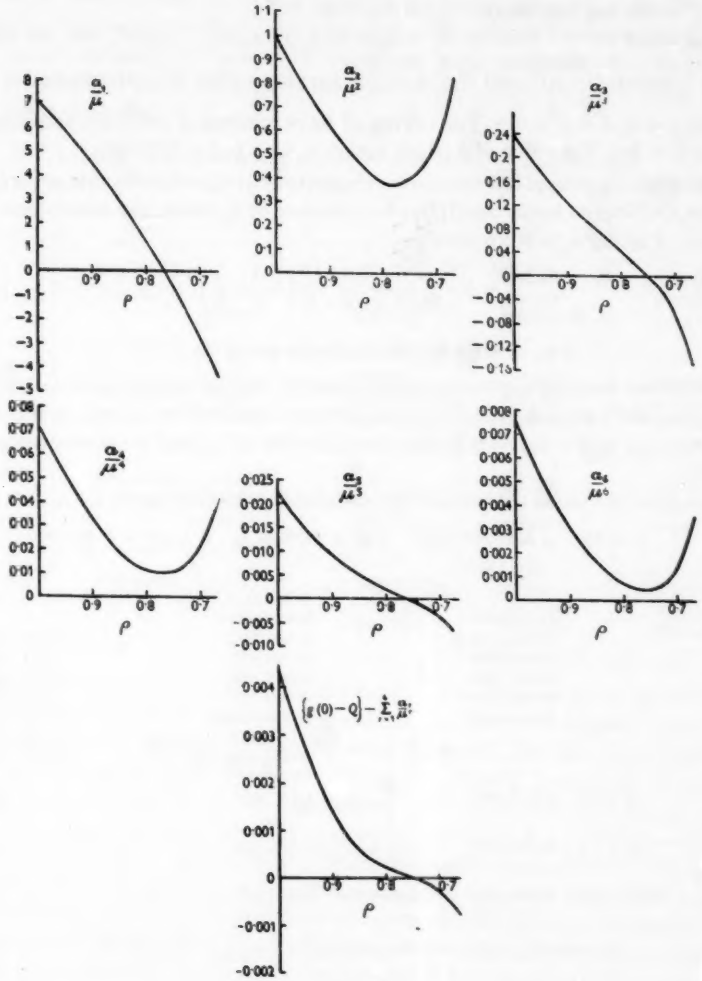

Fig. 1

2·3. *Choice of the value of* ρ

In order that the series should be of practical utility, it must be possible to represent $\exp\{\Sigma\alpha_r\mu^{-r}(1-2it)^{-r}\}$ adequately by a reasonable number of terms of the exponential series; this can be done only if the coefficients α are fairly small. In the *univariate* case, these coefficients will be small even if $\rho = 1$, and in fact if we put $p = 1$ and $\rho = 1$ in equation (26) the series we obtain corresponds exactly with that found by Hartley (1940) using rather a different method of approach. The accuracy of Hartley's series, using only three terms in the asymptotic series, was demonstrated (Hartley & Pearson, 1946) by comparison with the significance levels obtained from Nair's exact expansion; the agreement obtained was good even when the degrees of freedom were as low as three. In the multivariate case, a much more satisfactory series can be obtained if ρ is less than unity.

A typical set of curves showing the values of α_1/μ, α_2/μ^2, ..., α_6/μ^6, and the closeness of agreement between $Q-g(0)$ and $-\sum_{r=1}^{6}\alpha_r/\mu^r$ for varying values of ρ, are plotted in the figure for the case $p = 5$, $k = 5$, $\nu = 9$. The curves all have minima or cross the zero line between $\rho = 0{\cdot}7$ and $\rho = 0{\cdot}8$. The value of ρ which makes α_1 zero is $\rho = 0{\cdot}76296$.

In the calculations carried out here, ρ was chosen so that $\alpha_1 = 0$, since this not only resulted in the other coefficients being small, but the absence of α_1 made the calculation of the a's much easier. Putting $\alpha_1 = 0$ we obtain

$$\rho = 1-\frac{(2p^2+3p-1)}{6(p+1)(k-1)}\left(\Sigma\frac{1}{\nu_l}-\frac{1}{N}\right). \tag{44}$$

2·4. *Example of a calculation using the series*

To check his two working approximations (*a*) and (*b*), Bishop used as a standard of reference the values obtained by exact fitting of type I curves to the first two moments of the criterion l_1. In the case $p = 4$, $k = 5$, $\nu = 9$ Bishop found for the 5 % point a value corresponding to $M = 70{\cdot}281$.

To obtain from the series the probability associated with this value, we calculate

$$f = 40, \quad \rho = 0{\cdot}808,889, \quad \rho M = 56{\cdot}849,5, \quad \mu = \rho\nu = 7{\cdot}28.$$

$r\ (=v)$	α_r/μ^r	a_v/μ^v	P_{f+2v}
0	—	1·000,000	0·040,742
1	0·000,000	0·000,000	—
2	0·143,702	0·143,702	0·092,597
3	0·003,675	0·003,675	0·131,138
4	0·001,793	0·012,118	0·178,763
5	0·000,094	0·000,622	0·235,161
6	0·000,032	0·000,791	0·304,909
7	—	0·000,059	0·369,563
8	—	0·000,044	0·446,178
$\sum_1^6 \alpha_r/\mu^r$	0·149,296	$\sum_1^8 a_v/\mu^v$	1·161,011
$g(0)-Q$	0·149,305	$\exp\{\sum_1^6 \alpha_r/\mu^r\}$	1·161,016
Difference	0·000,009	Difference	0·000,005

$$K = \exp\{Q-g(0)\} = 0{\cdot}861,306, \quad \exp\left\{-\sum_1^6\alpha_r\Big/\mu^r\right\} = 0{\cdot}861,314$$

$$\Pr.\{M > 70{\cdot}281\} = K\Sigma\{a_v/\mu^v\}P_{f+2v} = 0{\cdot}0492.$$

To illustrate the accuracy with which the asymptotic series represents the function, independent calculations of $g(0)-Q$ have been made. As has already been indicated, however, in practice this rather laborious calculation would not be necessary, K being taken as

$$\exp\{-\Sigma\alpha_r/\mu^r\}.$$

2·5. *Some comparisons between the series and the exact distribution*

For the cases $p = 1, k = 2$ and $p = 2, k = 2$, the exact distribution is known for all values of ν; for $p = 1$ the criterion will simply be a function of the variance ratio, and when $p = 2$ the exact distribution has been found by Pearson & Wilks (1933). Table 1 enables the probabilities obtained from these exact distributions to be compared with those found using the series with scale factor ρ and up to four terms in the asymptotic and exponential series, higher terms having negligible effect. The table shows the values of M corresponding to the 5% and 1% points obtained by Bishop by fitting a type I curve to the first two moments of l_1. The exact probabilities corresponding to these points and those obtained using the series are shown below the values of M.

Table 1. *Comparison of the series with the exact distribution*

		$\nu=9$	$\nu=27$	$\nu=79$
$p=1$ $k=2$	5% point (type I)	4·0499	3·9042	3·8794
	Probability: exact	0·05005	0·05009	0·05002
	series	0·05005	0·05009	0·05003
	1% point (type I)	6·9902	6·7461	6·6991
	Probability: exact	0·00998	0·01001	0·01002
	series	0·00998	0·01001	0·01002
$p=2$ $k=2$	5% point (type I)	8·8801	8·1191	8·0018
	Probability: exact	0·05005	0·04997	0·04979
	series	0·05005	0·04997	0·04979
	1% point (type I)	12·8969	11·7844	11·6074
	Probability: exact	0·00999	0·01000	0·00997
	series	0·00999	0·01000	0·00997

The agreement between the series and the exact values is remarkably good, the series giving five-decimal accuracy in almost every case tested. The more difficult cases, however, are those where p and k are larger, especially when ν is small. For these, the closeness with which $\Sigma\alpha_r/\mu^r$ approaches $g(0)-Q$ and the adequacy, when ρ is suitably chosen, of the exponential series as judged by the comparison of $\exp\{\Sigma(\alpha_r/\mu^r)\}$ and $\Sigma(a_c/\mu^r)$, support belief in the accuracy of this solution. For example, the case $p = 4, k = 5$ which we have used to illustrate the calculation of probabilities from the series, is not a particularly favourable one. It appears, however, that six terms of the asymptotic series and eight of the exponential series will be adequate; in less severe cases of course fewer terms are necessary. Further evidence is supplied later for the accuracy of this type of solution, for in tests of independence to be discussed in § 6, exact distributions are available for comparison, in cases where the series is not favoured, and excellent agreement is found.

APPROXIMATIONS

The series we have found is of rather too complicated a character for routine use; as an alternative, approximations were sought which were relatively simple.

3. Approximations using a single χ^2 distribution

We have for the cumulant generating function of M (putting $\rho = 1$ in equation (20))

$$\Psi(t) = Q - g(0) - \frac{f}{2}\log(1-2it) + \sum_{r=1}^{\infty}\frac{\alpha'_r}{\nu^r}(1-2it)^{-r}, \tag{45}$$

where $f = \frac{1}{2}p(p+1)(k-1)$ and α'_r is obtained by putting $\rho = 1$ (i.e. $\beta = 0$) in equation (43).

Expanding this expression in powers of t we obtain

$$\Psi(t) = \sum_{j=1}^{\infty}\frac{(it)^j}{j!}2^{j-1}(j-1)!f\left\{1+\sum_{r=1}^{\infty}\binom{j+r-1}{r}\frac{2r\alpha'_r}{\nu^r f}\right\}. \tag{46}$$

The jth cumulant of M is then given by

$$\kappa_j = 2^{j-1}(j-1)!f\left\{1+jA_1+\frac{j(j+1)}{2}A_2+\ldots\right\}, \tag{47}$$

where

$$A_r = \frac{2r\alpha'_r}{\nu^r f}, \tag{48}$$

and in particular for the generalized test for homoscedasticity which we are considering,

$$\left.\begin{aligned} A_1 &= \frac{2p^2+3p-1}{6(k-1)(p+1)}\left(\sum_{l=1}^{k}\frac{1}{\nu_l}-\frac{1}{N}\right),\\ A_2 &= \frac{(p-1)(p+2)}{6(k-1)}\left(\sum_{l=1}^{k}\frac{1}{\nu_l^2}-\frac{1}{N^2}\right). \end{aligned}\right\} \tag{49}$$

3·1. *The choice of a scale factor in the χ^2 approximation*

Now $2^{j-1}(j-1)!f$ is the jth cumulant of χ^2 with f degrees of freedom. Thus, to order ν^{-1}, (47) is identical with the jth cumulant of $C\chi^2$, where C is either $1+A_1$ or $(1-A_1)^{-1}$. If A_2 were zero then $C = 1+A_1$ would give the first cumulant κ_1 to order ν^{-2} and the remaining cumulants would clearly be less in error than if C were taken as $(1-A_1)^{-1}$. However, if $A_2 = A_1^2$, it would be preferable to put $C = (1-A_1)^{-1}$, since here this form would give agreement to order ν^{-2}.

Clearly this would also be the better form to use if A_2 were near to or greater than A_1^2. In the univariate case $A_2 = 0$ and C should therefore be taken as

$$C = 1+A_1 = 1+\frac{1}{3(k-1)}\left(\Sigma\frac{1}{\nu_l}-\frac{1}{N}\right)$$

as has been shown by Bartlett (1937).

For the generalized test for homoscedasticity we find

$$A_2 - A_1^2 = \left(\frac{k}{k-1}\right)^2\frac{\gamma_2^2}{36(p+1)^2\nu^2}\left\{6(p-1)(p+1)^2(p+2)\left(\frac{\gamma_3}{\gamma_2^2}\frac{k-1}{k}\right)-(2p^2+3p-1)\right\}, \tag{50}$$

where γ_s is defined by equation (40). For $p = 1$, $A_2 = 0$, and consequently this quantity is negative for all values of k. When $p > 1$ it is positive, except in the particular case when $p = 2$ and $k = 2$ and the ν's are equal, when the quantity in curled brackets is equal to -1, and A_1^2 is almost exactly equal to A_2; if the ν's are not equal, this quantity is greater than -1, and it is positive for all larger values of p and k.

For the multivariate statistic, $p > 1$; we therefore take M/C to be approximately distributed as χ^2 with $f = \frac{1}{2}(k-1)p(p+1)$ degrees of freedom and

$$\frac{1}{C} = (1 - A_1) = 1 - \frac{(2p^2+3p-1)}{6(p+1)(k-1)}\left(\sum_{l=1}^{k}\frac{1}{\nu_l} - \frac{1}{N}\right);$$

if the degrees of freedom are equal this becomes

$$\frac{1}{C} = 1 - \frac{(2p^2+3p-1)(k+1)}{6(p+1)k\nu}. \tag{51}$$

We note that $1/C$ is the same as the value ρ chosen as scale factor (44) in the series solution. In the case of samples with equal degrees of freedom, the statistic M is equivalent to Bishop's criterion l_1, so that the multivariate scale factor C proposed here is comparable with the scale factor G proposed by Bishop and given in equation (3). Table 2 shows a number of comparisons for the significance levels, together with the values for the probabilities given by the series.

Table 2. *χ^2 approximation; comparisons of scale factors. Significance points for M with probability given by series*

			$p=2$	$p=4$	$p=6$
$k=5$ $\nu=9$	5%	Bishop (*b*)	23·06 0·0531	67·38 0·0742	142·19 0·1633
		Box	23·27 0·0503	68·93 0·0597	148·30 0·1041
	1%	Bishop (*b*)	28·82 0·0107	77·01 0·0173	156·35 0·0533
		Box	29·01 0·0101	78·74 0·0135	163·16 0·0286
$k=5$ $\nu=19$	5%	Bishop (*b*)	22·13 0·0486	61·36 0·0511	121·29 0·0660
		Box	22·03 0·0501	61·31 0·0515	122·82 0·0556
	1%	Bishop (*b*)	27·34 0·0104	69·95 0·0106	133·60 0·0144
		Box	27·47 0·0100	70·03 0·0105	135·13 0·0116

It appears that, not only is the factor suggested here very much simpler than Bishop's, but that it also gives a better approximation. However, it appears that even with the scale factor C this approximation fails when p is large and ν is small.

4. Approximations using the F distribution

The χ^2 approximation becomes less and less satisfactory as p and k are made larger and ν is made smaller. We know, however, that for all finite p and k, M/C will tend to a type III curve as ν becomes large. When ν is not large we might expect the point corresponding to the

distribution of M in the β_1, β_2 plane to lie near the type III line, in either the type I or type VI regions. We shall see that the use of these curves rather than the type III will enable us to absorb a further term in the cumulant series, corresponding to the extra adjustable parameter available with type I and type VI curves, and thus ensure agreement in the cumulants to order ν^{-2}. Although percentage points of the B-function have been tabled (Thompson, 1941), tables of the function F are usually more readily available. For this reason results which occur in the B-function form will be inverted, so that only tables of the F distribution will be required in using these approximations, and they will be referred to as F approximations.

4·1. *Choice of relevant type of curve*

The 'start' of the probability density function for M is at zero. For the Pearson system of frequency curves in which the restriction is made that the start of the curve is at zero, the relation between the cumulants

$$\frac{\kappa_3\kappa_1}{\kappa_2^2} = 2\tau \tag{52}$$

corresponds with Pearson's type III curve when $\tau = 1$. If τ slightly exceeds unity the curve falls in the type VI region, if it is slightly less than unity it falls in the type I region. Substituting the values for the cumulants of the criterion M, using equation (47), we obtain, ignoring terms of order ν^{-3},

$$\tau = \frac{1+4A_1+7A_2+3A_1^2}{1+4A_1+6A_2+4A_1^2}. \tag{53}$$

Thus for all sufficiently large values of ν the region into which the curve will fall is given by

$$\left.\begin{array}{ccc} A_2 > A_1^2 & A_2 = A_1^2 & A_2 < A_1^2 \\ \tau > 1 & \tau = 1 & \tau < 1 \\ \text{Type VI} & \text{Type III} & \text{Type I} \end{array}\right\} \tag{54}$$

For example, from equation (50) obtained in the case of the generalized test for homoscedasticity, it is clear that for $p = 1$ the curve will be in the type I region, and for nearly all other cases, when p is greater than 1, it will be in the type VI region.

4·2. *Type VI*

The F distribution with $2P$ and $2Q$ degrees of freedom is defined by

$$p(F) = \text{constant } F^{P-1}(PF+Q)^{-(P+Q)}. \tag{55}$$

The rth moment of a quantity bF, where b is a constant, is given by

$$\mu_r'(bF) = \left(b\frac{Q}{P}\right)^r \frac{\Gamma(P+r)\,\Gamma(Q-r)}{\Gamma(P)\Gamma(Q)}, \tag{56}$$

from which, after some algebraic reduction, we obtain the first four cumulants of bF as

$$\begin{aligned}
\kappa_1(bF) &= P(b/P)\,(1-1/Q)^{-1}, \\
\kappa_2(bF) &= P(b/P)^2\,(1+(P-1)/Q)\,(1-1/Q)^{-2}\,(1-2/Q)^{-1}, \\
\kappa_3(bF) &= 2P(b/P)^3\,(1+(P-1)/Q)\,(1+(2P-1)/Q)\,(1-1/Q)^{-3}\,(1-2/Q)^{-1}\,(1-3/Q)^{-1}, \\
\kappa_4(bF) &= 6P(b/P)^4\,\{(P/Q)^2\,(5/Q-11/Q^2)+(1-1/Q)^2\,(1-3/Q+2/Q^2+6P/Q-13P/Q^2)\} \\
&\quad \times (1-1/Q)^{-4}\,(1-2/Q)^{-2}\,(1-3/Q)^{-1}\,(1-4/Q)^{-1}.
\end{aligned} \tag{57}$$

Now we have seen that M is approximately distributed as $C\chi^2$, so that if τ is greater than unity we would expect to be able to find values b, P and Q, so that bF would be an even better approximation. Since we already know that the distribution is close to type III, we would further expect that Q will be large compared with P since this will be so for type VI curves close to the type III line.

If then we ignore terms of order $(P/Q)^2$, we find

$$\left.\begin{aligned}\kappa_1(bF) &= P(b/P)\{1 + 1/Q\},\\ \kappa_2(bF) &= P(b/P)^2\{1 + P/Q + 3/Q\},\\ \kappa_3(bF) &= 2P(b/P)^3\{1 + 3P/Q + 6/Q\},\\ \kappa_4(bF) &= 6P(b/P)^4\{1 + 6P/Q + 10/Q\}.\end{aligned}\right\} \tag{58}$$

Now put $2P = f_1 = f$, $2Q = f_2 = \dfrac{f_1+2}{A_2 - A_1^2}$ and $b = \dfrac{f_1}{1 - A_1 - f_1/f_2}$,

then we obtain approximately

$$\left.\begin{aligned}\kappa_1(bF) &= f\{1 + A_1 + A_2\},\\ \kappa_2(bF) &= 2f\{1 + 2A_1 + 3A_2\},\\ \kappa_3(bF) &= 8f\{1 + 3A_1 + 6A_2\},\\ \kappa_4(bF) &= 48f\{1 + 4A_1 + 10A_2\},\end{aligned}\right\} \tag{59}$$

which are identical to order ν^{-2} with the cumulants of M given by equation (47). Thus M/b will be distributed approximately as F with f_1 and f_2 degrees of freedom, where

$$f_1 = f, \quad f_2 = \frac{f_1+2}{A_2 - A_1^2}, \quad b = \frac{f_1}{1 - A_1 - f_1/f_2}. \tag{60}$$

4·3. *Type I*

We define a quantity X distributed in a type I form with parameters P and Q,

$$p(X) = \text{constant}\, X^{P-1}(1-X)^{Q-1}. \tag{61}$$

The rth moment of bX, where b is a constant, is given by

$$\mu'_r(bX) = b^r \frac{\Gamma(P+r)\,\Gamma(P+Q)}{\Gamma(P)\,\Gamma(P+Q+r)}, \tag{62}$$

from which we find the first four cumulants of bX to be

$$\left.\begin{aligned}\kappa_1(bX) &= P(b/Q)(1 + P/Q)^{-1},\\ \kappa_2(bX) &= P(b/Q)^2(1 + P/Q)^{-2}(1 + (P+1)/Q)^{-1},\\ \kappa_3(bX) &= 2P(b/Q)^3(1 - P/Q)(1 + P/Q)^{-3}(1 + (P+1)/Q)^{-1}(1 + (P+2)/Q)^{-1},\\ \kappa_4(bX) &= 6P(b/Q)^4\{1 + 1/Q - 2P/Q - 4P/Q^2 - 2P^2/Q^2 + P^2/Q^3 + P^3/Q^3\}\\ &\quad \times (1 + P/Q)^{-4}(1 + (P+1)/Q)^{-2}(1 + (P+2)/Q)^{-1}(1 + (P+3)/Q)^{-1}.\end{aligned}\right\} \tag{63}$$

As before if Q is large compared with P, so that terms of order $(P/Q)^2$ may be ignored, we obtain

$$\left.\begin{aligned}\kappa_1(bX) &= P(b/Q)\{1 - P/Q\},\\ \kappa_2(bX) &= P(b/Q)^2\{1 - 3P/Q + 1/Q\},\\ \kappa_3(bX) &= 2P(b/Q)^3\{1 - 6P/Q - 3/Q\},\\ \kappa_4(bX) &= 6P(b/Q)^4\{1 - 10P/Q - 6/Q\},\end{aligned}\right\} \tag{64}$$

and putting $2P = f_1 = f$, $2Q = f_2 = \dfrac{f_1+2}{A_1^2 - A_2}$ and $b = \dfrac{f_2}{1 - A_1 + 2/f_2}$

we again obtain approximately the values given in (59) which to the order of approximation ν^{-2} are the cumulants of M.

Thus M/b will be distributed as X in expression (61) with $2P = f_1$ and $2Q = f_2$ and

$$f_1 = f, \quad f_2 = \frac{f_1+2}{A_1^2 - A_2}, \quad b = \frac{f_2}{1 - A_1 + 2/f_2}. \tag{65}$$

Alternatively, $\dfrac{f_2 M}{f_1(b-M)}$ will be distributed as F with f_1 and f_2 degrees of freedom.

We note that although M can vary from 0 to ∞, bX can vary only between the limits 0 and b, so that we are fitting a curve with limited range to one with infinite range. In practice, however, this presents no difficulty (see, for example, the comparisons of Tables 3, 4 and 5), for since the distribution of M will be near to type III, f_2 will be large compared with f_1; consequently b will be large compared with f_1. The mean for such curves will be approximately equal to f_1, so that the range will be large compared with the mean, and the part of the curve ignored by the truncation will be negligible.

4·4. *Application of the F approximation in tests of homoscedasticity*

From (50) we know that when $p = 1$, $A_2 - A_1^2$ is negative, and hence the type I form of the approximation is appropriate. When $p \geqslant 2$ we have seen that, except for the case $p = 2, k = 2$, when to this degree of approximation the curve is almost type III, $A_2 - A_1^2$ is positive and the type VI form is appropriate.

4·41. *Univariate test* ($p = 1$)

When $p = 1$, A_2 is zero, so that to carry out the test we calculate in turn

$$A_1 = \frac{1}{3(k-1)}\left(\Sigma\frac{1}{\nu_i} - \frac{1}{N}\right), \quad f_1 = (k-1), \quad f_2 = \frac{k+1}{A_1^2}, \quad b = \frac{f_2}{1 - A_1 + 2/f_2}, \tag{66}$$

and refer $\dfrac{f_2 M}{f_1(b-M)}$ to tables of the F distribution with f_1 and f_2 degrees of freedom.

In the special case when the degrees of freedom are equal

$$A_1 = \frac{k+1}{3\nu k}. \tag{67}$$

To test the accuracy of the approximation we will compare the values it gives for the 5% and 1% points of M, with those obtained from (1) Bartlett's approximation, (2) Bishop & Nair's (1939) values and (3) the χ^2 series given by Hartley and corresponding to equation (26) with $p = 1$ and $\rho = 1$. Tables 3 and 4 are adapted from those given by Pearson & Hartley (1946) with the value of Bartlett's approximation and the present approximation added. In Table 3 a number of comparisons are made for the special case where the degrees of freedom are equal, and Table 4 shows a few comparisons for the case of five estimates of variance with unequal degrees of freedom.

If the accuracy is judged by the closeness of agreement with the values obtained by Bishop & Nair, it appears that the F approximation is an improvement upon that suggested by

Table 3. *Comparison of approximations. Significance points for M (equal degrees of freedom, $p = 1$)*

k	ν	5%				1%			
		Bartlett (χ^2)	Box (F)	Hartley (series)	Bishop & Nair	Bartlett (χ^2)	Box (F)	Hartley (series)	Bishop & Nair
3	2	7·32	7·20	7·05	7·11*	11·26	10·85	10·57	10·74*
	3	6·88	6·83	6·79	6·80†	10·57	10·41	10·32	10·43†
	4	6·66	6·63	6·61	6·62*	10·23	10·14	10·10	10·13*
	9	6·29	6·29	6·28	6·30†	9·67	9·64	9·64	9·67†
5	2	11·39	11·23	11·01	11·09*	15·93	15·52	15·15	15·32*
	3	10·75	10·69	10·62	10·67†	15·05	14·88	14·76	14·91†
	4	10·44	10·39	10·37	10·38*	14·60	14·42	14·46	14·47*
	9	9·91	9·89	9·90	9·93†	13·87	13·85	13·84	13·86†
10	2	20·02	19·68	19·45	19·62*	25·68	25·22	24·65	24·90*
	3	18·99	18·91	18·79	18·82†	24·31	24·12	23·97	24·09†
	4	18·47	18·41	18·38	18·42*	23·65	23·54	23·49	23·34*
	9	17·61	17·61	17·60	17·64†	22·69	22·58	22·53	22·48†

* Calculated from Nair's exact distribution. † Calculated by fitting type I curve to L_1.

Table 4. *Comparison of approximations. Significance points for M (unequal degrees of freedom, $p = 1$, $k = 5$)*

N	ν_1	ν_2	ν_3	ν_4	ν_5	5%				1%			
						Bartlett (χ^2)	Box (F)	Hartley (series)	Bishop & Nair	Bartlett (χ^2)	Box (F)	Hartley (series)	Bishop & Nair
20	6	6	4	2	2	10·70	10·65	10·54	10·59	14·97	14·82	14·62	14·80
45	16	16	9	2	2	10·45	10·40	10·30	10·35	14·62	14·51	14·31	14·46
20	5	5	4	3	3	10·49	10·46	10·41	10·43	14·68	14·58	14·51	14·59
45	14	14	9	4	4	10·07	10·05	10·04	10·05	14·09	14·05	14·03	14·05

Bartlett and is about as accurate as Hartley's series, whilst it requires no special tables and involves only simple calculations.

Since the approximations proposed by Bartlett, Hartley and the present author are essentially asymptotic, it is to be expected that for small values of ν, and particularly when $\nu = 1$, the approximations will break down. This does in fact happen to a certain extent with all of them, but it seems least serious with the present F approximation; for example, when $k = 4$, $\nu = 1$, we have

Approximation	5% point	1% point
Bartlett (χ^2)	11·1	16·1
Hartley (series)	9·0	11·8
Box (F)	10·3	14·6
Nair's expansion	10·0	14·1

For the case $\nu = 1$, Table 5 compares, for a number of values of k, the 5% and 1% levels given by Bartlett's approximation and by the present method with values obtained by Bishop & Nair (1939) using Nair's expansion.

Table 5. *Comparison of the approximations when* $\nu = 1$. *Significance points for* M

Value of k		2	3	4	5	6	7	8	9	10
5% point	Bartlett (χ^2)	5·8	8·7	11·1	13·3	15·4	17·4	19·3	21·3	23·1
	Box (F)	5·1	7·9	10·3	12·6	14·6	16·7	18·6	20·5	22·4
	Nair's expansion	5·1	7·7	10·0	12·0	14·1	15·9	17·9	19·6	21·3
1% point	Bartlett (χ^2)	10·0	13·3	16·1	18·6	21·0	23·2	25·4	27·5	29·6
	Box (F)	7·9	11·3	14·6	17·1	19·2	21·5	23·7	25·8	27·9
	Nair's expansion	8·3	11·5	14·0	16·5	18·9	21·0	23·1	25·2	27·2

4·42. *Multivariate test* $p \geqslant 2$

To carry out the test we calculate the quantities

$$\left.\begin{aligned} A_1 &= \frac{2p^2+3p-1}{6(k-1)(p+1)}\left(\Sigma\frac{1}{\nu_i}-\frac{1}{N}\right), \quad A_2 = \frac{(p-1)(p+2)}{6(k-1)}\left(\Sigma\frac{1}{\nu_i^2}-\frac{1}{N^2}\right), \\ f_1 &= \tfrac{1}{2}(k-1)p(p+1), \quad f_2 = \frac{f_1+2}{A_2-A_1^2}, \quad b = \frac{f_1}{1-A_1-f_1/f_2}, \end{aligned}\right\} \tag{68}$$

and refer M/b to the tables of the F distribution with f_1 and f_2 degrees of freedom.

When the degrees of freedom are equal

$$A_1 = \frac{(p^2+3p-1)(k+1)}{6(p+1)k\nu}, \quad A_2 = (p-1)(p+2)\frac{(k^2+k+1)}{6k^2\nu^2}. \tag{69}$$

George (1945) was able to evaluate the exact distribution of the generalized L_1 statistic in simple cases, although, when the value of p and k are not very small, the method becomes unmanageable. She used her exact distribution to check Bishop's approximations. Table 6 is taken for George's Table 1 and shows the equivalent value of M obtained by Bishop's empirical formula, method (a), for the 5% point, together with the exact value of the probability obtained by George by direct integration. The probability corresponding to this value of M has also been calculated by the χ^2 and F approximations suggested here. Thus the closeness with which exact probability approaches 0·0500 indicates the accuracy of Bishop's method, and the closeness with which the probabilities for the χ^2 and F approximations coincide with the exact probability measures the accuracy of these approximations.

We see that the values given by the F approximation are in excellent agreement with the exact probabilities, and even the χ^2 approximation is considerably better than Bishop's method. Unfortunately, no exact values are available in the cases where p and k are larger, when approximation to the curve is more difficult. For these distributions the series given by formula (26), using in most cases up to six* terms in the asymptotic series and up to eight*

* When $\nu = 9$, and $p = 5$ and 6, the coefficients α are rather large, and ten and fourteen terms respectively had to be used in the exponential series. When $p = 6$ there is evidence that further terms in the asymptotic series would give closer agreement.

Table 6. *5 % points for M given by Bishop's empirical approximation with their associated probabilities calculated by:* (1) *George's exact method,* (2) *the F approximation,* (3) *the* χ^2 *approximation*

	ν	M	Probabilities				ν	M	Probabilities		
			Exact (George)	F (Box)	χ^2 (Box)				Exact (George)	F (Box)	χ^2 (Box)
$p=2$	9	8·924	0·0492	0·0492	0·0492	$p=3$	9	15·740	0·0461	0·0458	0·0446
$k=2$	14	7·835	0·0495*	0·0494	0·0496	$k=2$	14	14·434	0·0475	0·0475	0·0470
	19	7·831	0·0496*	0·0496	0·0496		24	13·598	0·0485	0·0486	0·0485
	24	7·828	0·0498*	0·0497	0·0497		29	13·416	0·0488	0·0489	0·0487
							39	13·211	0·0488	0·0489	0·0488
$p=2$	9	14·164	0·0491	0·0491	0·0490						
$k=3$	19	13·285	0·0496	0·0497	0·0496	$p=3$	14	23·661	0·0481	0·0479	0·0473
	29	13·031	0·0497	0·0499	0·0499	$k=3$	29	22·288	0·0484	0·0480	0·0478
						$p=4$	19	20·989	0·0461	0·0461	0·0455
						$k=2$	29	19·946	0·0477	0·0478	0·0476

* These values have been recalculated and do not agree with the values given in George's table.

Table 7. *Comparisons of approximations. Significance points for M obtained by various methods, with probabilities given by series* (26)

			$p=2$		$p=3$		$p=4$		$p=5$		$p=6$	
$k=5$	5%	Bishop (*a*)	23·40	0·0485	—	—	71·07	0·0434	—	—	173·17	0·0105
$\nu=9$		Bishop (*b*)	23·06	0·0531	—	—	67·38	0·0742	—	—	142·19	0·1633
		Box (χ^2)	23·27	0·0503	42·56	0·0532	68·93	0·0597	103·65	0·0673	148·30	0·1041
		Box (F)	23·30	0·0500	42·83	0·0506	69·84	0·0524	106·40	0·0545	153·36	0·0692
		Type I	23·26	0·0504	42·88	0·0502	70·28	0·0492	107·15	0·0500	157·38	0·0488
	1%	Bishop (*a*)	29·19	0·0096	—	—	81·35	0·0082	—	—	192·34	0·0010
		Bishop (*b*)	28·82	0·0107	—	—	77·01	0·0173	—	—	156·35	0·0533
		Box (χ^2)	29·01	0·0101	50·24	0·0111	78·74	0·0135	113·54	0·0225	163·16	0·0286
		Box (F)	29·05	0·0100	50·58	0·0102	79·84	0·0105	118·56	0·0122	168·94	0·0165
		Type I	29·07	0·0099	50·59	0·0102	80·45	0·0097	120·20	0·0098	173·78	0·0097
$k=5$	5%	Bishop (*a*)	22·14	0·0486	—	—	61·63	0·0489	—	—	124·25	0·0469
$\nu=19$		Bishop (*b*)	22·13	0·0486	—	—	61·36	0·0511	—	—	121·29	0·0660
		Box (χ^2)	22·03	0·0501	39·09	0·0506	61·31	0·0515	89·08	0·0532	122·82	0·0556
		Box (F)	22·04	0·0500	39·14	0·0501	61·47	0·0502	89·47	0·0505	123·61	0·0508
		Type I	21·92	0·0516	39·10	0·0505	61·36	0·0511	89·59	0·0496	123·70	0·0503
	1%	Bishop (*a*)	27·55	0·0098	—	—	70·50	0·0095	—	—	137·09	0·0088
		Bishop (*b*)	27·34	0·0104	—	—	69·95	0·0106	—	—	133·60	0·0144
		Box (χ^2)	27·47	0·0100	46·14	0·0102	70·03	0·0105	99·56	0·0109	135·13	0·0116
		Box (F)	27·48	0·0100	46·20	0·0100	70·23	0·0101	100·01	0·0101	136·03	0·0103
		Type I	27·55	0·0098	46·24	0·0099	70·27	0·0100	100·24	0·0098	136·32	0·0099

terms in the exponential series, may be used as a standard for comparison. Table 7 shows the significance points for M obtained by five different methods together with the probabilities calculated from the series. The methods are: Bishop's empirical approximation (a), Bishop's approximation (b), the χ^2 and F approximations suggested in this paper, and the fitting of a type I curve by exact calculation of the first two moments. The values for M for Bishop's approximations and the type I approximation have been calculated from Bishop's significance points for l_1 given in his Tables 9 and 10.

If we take the series solution as supplying essentially accurate values, we confirm Bishop's suggestion that the type I curve, fitted exactly to the first two moments of l_1, provides an exceedingly good approximation. Of the working approximations, the F approximation suggested here appears to be the best and the χ^2 approximation with the generalized scale factor C will be fairly satisfactory if p and k are not greater than five and ν is not less than, say, twenty.

Table 8 supplies a few comparisons with equal and unequal degrees of freedom.

Table 8. *Significance points for M from χ^2 and F approximations for some equal and unequal groupings, when $p = 4$ and $k = 5$, with associated probability given by series* (26)

N	ν_1	ν_2	ν_3	ν_4	ν_5		5%		1%	
95	19	19	19	19	19	χ^2	61·31	0·0515	70·03	0·0105
						F	61·47	0·0502	70·23	0·0101
95	9	9	19	29	29	χ^2	63·22	0·0578	72·33	0·0124
						F	63·99	0·0521	73·14	0·0107
95	9	9	9	9	59	χ^2	66·32	0·0627	75·76	0·0139
						F	67·39	0·0535	77·07	0·0110
45	9	9	9	9	9	χ^2	68·93	0·0597	78·74	0·0135
						F	69·84	0·0524	79·84	0·0105

It appears, at least for unequal samples with none of the degrees of freedom less than 9, that the F approximation will be fairly satisfactory.

5. Generalization of the procedure

The method we have developed has so far been illustrated in the case of the univariate and multivariate tests of homoscedasticity; its application is, however, more general. In fact, the method can be used whenever, by choosing a suitable power of the original criterion, we can obtain a statistic W which has its hth moment of the form

$$\mathscr{E}(W)^h = \text{constant} \times \left[\frac{\prod_{j=1}^{k}(y_j^{y_j})}{\prod_{i=1}^{m}(x_i^{x_i})}\right]^h \frac{\prod_{i=1}^{m}[\Gamma\{x_i(1+h)+\xi_i\}]}{\prod_{j=1}^{k}[\Gamma\{y_j(1+h)+\eta_j\}]}, \tag{70}$$

where $$\sum_{i=1}^{m} x_i = \sum_{j=1}^{k} y_j.$$

(The constant will be obtained of course by putting $h = 0$ in (70) and taking the reciprocal.) Many of the tests in Plackett's review, referred to in the introduction to this paper, fall into this category. We have already seen that the generalized L_1 statistic is of this type; others are Wilks's test for the independence of k groups of variates (which has some important special cases; and will be considered in detail in the next section); the generalized test for constancy of means, variances and covariances for k samples given by Wilks (1932); and the tests for 'compound symmetry' of variance-covariance matrices discussed by Votaw (1948).

Another group of criteria, which has been studied by Mauchly (1940) and Wilks (1946), arises from tests made on a single sample of n p-variate observations. Mauchly's criterion tests the hypothesis that the variances of the variates are all equal and that the covariances between the variates are all zero. Wilks considered criteria for testing three further hypotheses:

(*a*) That the p means, p variances, and $\frac{1}{2}p(p-1)$ covariances for the variates have respectively the same unknown values.

(*b*) That the variances are the same and the covariances are the same irrespective of what values the means have.

(*c*) That the means are the same (assuming (*b*) true).

It is hoped to consider some of these tests rather more closely in a later paper. Here we shall merely note that, except for Wilks's third criterion (which is always distributed exactly in type I form), the exact distribution of the test function is, in general, not exactly known. The expression for the hth moment, however, is in each case of the form of equation (70) and, as is shown below, our previous approach will provide approximations in all these cases. Tukey & Wilks (1946) have considered this class of statistics and have pointed out that they all possess in common the property that, when the null hypothesis is true, they are distributed as a product of independent components, each component being distributed in type I form.

Consider the expression (70) for the hth moment of any statistic W of this type. If we take $M = -2\log W$ as our working statistic, and write $(1-\rho)x_i = \beta_i$, $(1-\rho)y_j = \epsilon_j$, where ρ is a constant $\leqslant 1$ at our choice, we find for the cumulant generating function of ρM

$$\Psi(t) = g(t) - g(0),$$

where
$$g(t) = 2it\rho\left[\sum_{i=1}^{m} x_i \log x_i - \sum_{j=1}^{k} y_j \log y_j\right] + \sum_{i=1}^{m} \log\Gamma\{\rho x_i(1-2it) + \beta_i + \xi_i\} - \sum_{j=1}^{k} \log\Gamma\{\rho y_j(1-2it) + \epsilon_j + \eta_j\}, \quad (71)$$

and $g(0)$ is independent of t and is obtained by writing $t = 0$ in (71). Expanding the logarithms of the Γ-functions by (18), we obtain the cumulant generating function of ρM in the form

$$\Psi(t) = Q - g(0) - \frac{f}{2}\log(1-2it) + \sum_{r=1}^{\infty} \omega_r (1-2it)^{-r}, \quad (72)$$

where
$$f = -2\left\{\sum_{i=1}^{m}\xi_i - \sum_{j=1}^{k}\eta_j - \tfrac{1}{2}(m-k)\right\}, \tag{73}$$

$$\omega_r = \frac{(-1)^{r+1}}{r(r+1)}\left\{\sum_{i=1}^{m}\frac{B_{r+1}(\beta_i+\xi_i)}{(\rho x_i)^r} - \sum_{j=1}^{k}\frac{B_{r+1}(\epsilon_j+\eta_j)}{(\rho y_j)^r}\right\}, \tag{74}$$

$$Q = \tfrac{1}{2}(m-k)\log 2\pi - \frac{f}{2}\log\rho + \sum_{i=1}^{m}(x_i+\xi_i-\tfrac{1}{2})\log x_i - \sum_{j=1}^{k}(y_j+\eta_j-\tfrac{1}{2})\log y_j. \tag{75}$$

From the cumulant generating function (72), the asymptotic χ^2 series corresponding to (26) and (30) are immediately obtainable. Alternatively, we may obtain approximations in the manner given in § 3; the method outlined there is clearly perfectly general for this whole class of statistics. We need the quantities $A_1 = 2\omega_1'/f$ and $A_2 = 4\omega_2'/f$, where ω_r' is the value taken on by ω_r when $\rho = 1$. The scale factor C for the χ^2 approximation will be $1+A_1$ or $(1-A_1)^{-1}$; a decision between the two alternative forms can be reached by the considerations set out in § 3·1. Then, to this order of approximation, M/C will be distributed as χ^2. If greater accuracy is required we may use the F type approximations described in § 4. The particular form is decided by the sign of the quantity $A_2 - A_1^2$. If this quantity is positive, the curve of best fit will be type VI. Putting

$$f_1 = f, \quad f_2 = \frac{f_1+2}{A_2-A_1^2}, \quad b = \frac{f_1}{1-A_1-f_1/f_2}, \tag{76}$$

M/b is distributed approximately as the variance ratio F with f_1 and f_2 degrees of freedom. Alternatively, if $A_2 - A_1^2$ is negative the best fitting curve will be type I, and if we put

$$f_1 = f, \quad f_2 = \frac{f_1+2}{A_1^2-A_2}, \quad b = \frac{f_2}{1-A_1+2/f_2}, \tag{77}$$

then approximately $\dfrac{f_2 M}{f_1(b-M)}$ will be distributed as F with f_1 and f_2 degrees of freedom.

There are thus a number of possible levels of approximation as measured by the order of agreement between the cumulants of the statistic and those of the fitted curve.

(1) Ignoring terms of order x_i^{-1}, y_j^{-1}, M is distributed as χ^2.

(2) Ignoring terms of order x_i^{-2}, y_j^{-2}; by a technique originally used by Bartlett and here generalized, a quantity C can be found such that M is distributed as $C\chi^2$.

(3) Ignoring terms of order x_i^{-3}, y_j^{-3}, a function of M can be obtained which is distributed as the variance ratio F.

(4) Finally, for very precise work and for checking other approximations, a χ^2 series solution may be used and here agreement with the cumulants (as represented by their asymptotic expansions) of the statistic can be obtained to as great an order as seems profitable.

In practice method (4) is sometimes rather long, although it has been found very accurate, but (3) involves very little labour and will often be sufficiently precise.

As a second example of the application of this technique we consider Wilks's generalized test of independence.

6. The generalized test for independence

Wilks (1935) considered the following problem: suppose we have a sample of $\nu+u$ observations for a kp variate normal population and we have some *a priori* reason for dividing the variates into k groups containing $p_1, \ldots, p_n, \ldots, p_l, \ldots, p_k$ variates (where $\sum_l p_l = kp$ and p is thus the average size of the groups and is not necessarily integer). It is required to test

the hypothesis that the k groups of residuals, obtained after fitting u independent constants to each of the variates, are mutually independent.

If $|c_{ij}|$ is the $kp \times kp$ determinant of sums of squares and products of residuals for the kp variates and $|c_{ij}|_l$ is the $p_l \times p_l$ determinant of sums of squares and products of residuals of the lth group, then the likelihood ratio criterion obtained by Wilks is

$$\lambda = \frac{|c_{ij}|}{\prod_{l=1}^{k} |c_{ij}|_l} = \frac{|r_{ij}|}{\prod_{l=1}^{k} |r_{ij}|_l}, \tag{78}$$

where $|r_{ij}|$ and $|r_{ij}|_l$ are the corresponding determinants of sample correlation coefficients having ν degrees of freedom. Wilks obtained the moments, and also, for special sets of values of k and p_l, the exact distribution of his criterion which generalizes a very large class of statistical tests. Problems in which there are more than two groups of variates, i.e. where $k > 2$, occur for example in educational research; we may have some prior reason for believing that a battery of, say, ten different tests applied to pupils may be divided up into a number of groups, each group concerned with some distinct ability, and may wish therefore to test the hypothesis that, when the means are eliminated, the selected groups are independent of each other.

When $k = 2$, we consider only two groups of variates containing p_1 in the first and p_2 in the second. Since the criterion and its distribution will be unaffected if the set of p_2 variates are fixed independent variables and the set of p_1 variates 'dependent' variables distributed in a p_1-variate normal distribution, the function is then appropriate for testing the general multivariate linear hypothesis (see, for example, Bartlett, 1934, 1938, 1947). If, in addition, $p_1 = 1$, then the likelihood criterion is $\lambda = 1 - R^2$, where R is the coefficient of multiple correlation between the single dependent variate and the p_2 independent variates. A second special case of Wilks's statistic which is of some interest, and is considered more fully later in this section, occurs when there is only one variate in each of the k groups. The statistic then supplies an overall test for independence between k variates. For the general statistic Wald & Brookner (1941), using a rather different technique from that of Wilks, were able to extend the catalogue of values of k and p_l for which the distribution of λ is exactly known in terms of elementary functions, to include all cases where at most one group contains an odd number of variates. These distributions, although exact, are rather complicated in character. As an alternative and to cover the remaining cases, these authors obtained a series solution and Rao (1948) modified this series in the important case where $k = 2$ to provide an improved test in problems of multivariate analysis. These series will later appear as special cases of that which we are now investigating.

6·1. *Derivation of the series*

The hth moment of λ is given by Wilks as

$$\prod_{l=1}^{k} \prod_{j=0}^{p_l-1} \left\{ \frac{\Gamma\left(\frac{\nu-j}{2}\right)}{\Gamma\left(\frac{\nu-j}{2}+h\right)} \right\} \prod_{m=0}^{kp-1} \left\{ \frac{\Gamma\left(\frac{\nu-m}{2}+h\right)}{\Gamma\left(\frac{\nu-m}{2}\right)} \right\}. \tag{79}$$

So that if we write $W = \lambda^{\frac{1}{2}\nu}$, the hth moment of W will be in the form given in equation (70); taking as our logarithmic statistic $M = -2 \log W$ we obtain

$$M = -\nu \log \lambda. \tag{80}$$

To obtain the series, we begin as before by defining the relationship between a quantity ρ (less than unity) and quantities μ and β by the equations $\rho = \mu/\nu$, $\nu = \mu + \beta$. It is also convenient to define a set of quantities

$$\Sigma_s = \left(\sum_l p_l\right)^s - \sum_l p_l^s, \tag{81}$$

which appear in the solution in much the same way as the quantities $\Sigma \frac{1}{\nu_l^s} - \frac{1}{N^s}$ appear in the tests for homoscedasticity. Then, as before, we obtain equation (72) for the cumulant generating function of the ρM, and the constants are available by direct substitution in (73), (74) and (75),

$$f = \tfrac{1}{2}\Sigma_2, \tag{82}$$

$$\omega_r = \frac{\alpha_r}{\mu^r} = \frac{(-2)^r}{r(r+1)\mu^r} \sum_{l=2}^{k} \sum_{j=0}^{p_l - 1} \left\{ B_{r+1}\left(\frac{\beta - j}{2}\right) - B_{r+1}\left(\frac{\beta - \sum_{n=1}^{l-1} p_n - j}{2}\right) \right\}, \tag{83}$$

$$Q = -\frac{f}{2}\log\frac{\mu}{2}. \tag{84}$$

The calculation of α_r from formula (83) would clearly be extremely laborious for all but small numbers of variates; we therefore seek an alternative simpler form. Using relations (32) and (35), we find

$$\alpha_r = \frac{(-1)^{r+1}}{r(r+1)(r+2)} \sum_{s=0}^{r+1} \binom{r+2}{s+1} 2^s \sum_{l=2}^{k} \left\{ \beta^{r+1-s} - \left(\beta - \sum_{n=1}^{l-1} p_n\right)^{r+1-s} \right\} \delta_s(p_l), \tag{85}$$

where

$$\delta_s(p_l) \doteqdot B_{s+1}\left(-\frac{B + p_l}{2}\right) - B_{s+1}\left(-\frac{B}{2}\right), \tag{86}$$

and the values taken by (86) when $p = 1, 2, \ldots, 7$, are given by putting $p = p_l$ in equation (42). Writing α_r' for the value which α_r has when $\rho = 1$, i.e. when $\beta = 0$, then substituting for $\delta_s(p_l)$ in (85) and summing, we obtain for the first six values of α_r':

$$\left.\begin{aligned}
\alpha_1' &= \tfrac{1}{24}\{2\Sigma_3 + 3\Sigma_2\},\\
\alpha_2' &= \tfrac{1}{48}\{\Sigma_4 + 2\Sigma_3 - \Sigma_2\},\\
\alpha_3' &= \tfrac{1}{720}\{6\Sigma_5 + 15\Sigma_4 - 10\Sigma_3 - 30\Sigma_2\},\\
\alpha_4' &= \tfrac{1}{480}\{2\Sigma_6 + 6\Sigma_5 - 5\Sigma_4 - 20\Sigma_3 + 3\Sigma_2\},\\
\alpha_5' &= \tfrac{1}{8640}\{2\Sigma_7 + 7\Sigma_6 - 7\Sigma_5 - 35\Sigma_4 + 7\Sigma_3 + 49\Sigma_2\},\\
\alpha_6' &= \tfrac{1}{20160}\{3\Sigma_8 + 12\Sigma_7 - 14\Sigma_6 - 84\Sigma_5 + 21\Sigma_4 + 196\Sigma_3 - 10\Sigma_2\},
\end{aligned}\right\} \tag{87}$$

where Σ_s is defined by equation (81), whence we have for the α's

$$\left.\begin{aligned}
\alpha_1 &= \alpha_1' - (f/2)\beta,\\
\alpha_2 &= \alpha_2' - \alpha_1'\beta + (f/4)\beta^2,\\
\alpha_3 &= \alpha_3' - 2\alpha_2'\beta + \alpha_1'\beta^2 - (f/6)\beta^3,\\
\alpha_4 &= \alpha_4' - 3\alpha_3'\beta + 3\alpha_2'\beta^2 - \alpha_1'\beta^3 + (f/8)\beta^4,\\
\alpha_5 &= \alpha_5' - 4\alpha_4'\beta + 6\alpha_3'\beta^2 - 4\alpha_2'\beta^3 + \alpha_1'\beta^4 - (f/10)\beta^5,\\
\alpha_6 &= \alpha_6' - 5\alpha_5'\beta + 10\alpha_4'\beta^2 - 10\alpha_3'\beta^3 + 5\alpha_2'\beta^4 - \alpha_1'\beta^5 + (f/12)\beta^6.
\end{aligned}\right\} \tag{88}$$

As before, from the cumulant generating function we obtain the series corresponding to (26) and (30), and if ρ is chosen so that $\alpha_1 = 0$, we have

$$\beta = 2\alpha_1'/f, \quad \rho = 1 - \frac{1}{12\nu f}(2\Sigma_3 + 3\Sigma_2). \tag{89}$$

Wald & Brookner (1941) derived a χ^2 series for this statistic by a different method from that used here; it is not difficult to show, however, that the series they obtained is equivalent to our series, but with $\rho = 1$. In this form the series is of little practical use for small, or even moderate values of ν because of the difficulty we have noted before of adequately approximating to $\exp \Sigma\alpha_r\mu^{-r}(1-2it)^{-r}$ by means of a series, unless α_r is small or μ is large. By introducing the factor ρ, the size of the coefficients α can be greatly reduced and the series be used even for fairly small values of ν. As an example, consider the case of three groups of variates with two variates in each grouping, $k = 3$, $p_1 = 2$, $p_2 = 2$, $p_3 = 2$, and suppose $\nu = 10$. The values of the coefficients α_r/μ^r are shown below when $\rho = 1$ and also when ρ takes on a value making α_1 zero. When $\rho = 1$, μ is of course equal to ν.

Values of α_r/μ^r

r	$\rho = 1$	$\rho = 0{\cdot}683$
1	1·900,00	0·000,00
2	0·335,00	0·073,17
3	0·086,33	0·003,71
4	0·026,88	0·001,78
5	0·009,40	0·000,23
6	0·003,55	0·000,07
Total	2·361,16	0·078,96
$g(0)-Q$	2·363,61	0·078,98

For the Wald & Brookner series, if ν is small, the coefficients are so large that in practice it would be impossible to represent the exponent adequately by a reasonably small number of terms of the exponential series; by suitably choosing ρ, however, the size of the coefficients are greatly reduced while the agreement between the sum of the terms and $g(0)-Q$ is improved. In the particular example quoted, the exact distribution is known (Wilks, 1935). It appears in rather a complicated form, but has been used here to check the series and the approximations. Table 9 shows the 5% and 1% significance points for the criterion M

Table 9. *Some comparisons for Wilks's statistic*

			χ^2 approximation			F approximation		
			M	Probability		M	Probability	
				Exact	Series		Exact	Series
$k=3$ $p_1=2$ $p_2=2$ $p_3=2$	$\nu=10$	5%	30·770	0·0612	0·0612	31·357	0·0549	0·0548
		1%	38·366	0·0139	0·0139	39·180	0·0117	0·0117
	$\nu=20$	5%	24·982	0·0516	0·0516	25·083	0·0504	0·0504
		1%	31·149	0·0105	0·0105	31·292	0·0101	0·0101

obtained by using the χ^2 and F approximations which are derived in the next section, together with the exact probabilities and the probabilities calculated from the series, using $\rho = 0{\cdot}68\dot{3}$, and six terms in the asymptotic and eight in the exponential series. Agreement to four places of decimals is usually obtained between the series and the exact value for the probability.

χ^2 approximation. Following the previous procedure, we find that M/C is distributed approximately as χ^2 with f degrees of freedom, where

$$\frac{1}{C} = 1 - \frac{1}{12\nu f}(2\Sigma_3 + 3\Sigma_2) \quad \text{and} \quad f = \tfrac{1}{2}\Sigma_2.$$

F approximation. We have

$$f = \tfrac{1}{2}\Sigma_2, \quad A_1 = \frac{1}{12\nu f}(2\Sigma_3 + 3\Sigma_2), \quad A_2 = \frac{1}{12\nu^2 f}(\Sigma_4 + 2\Sigma_3 - \Sigma_2),$$

from which, using equations (76) and (77), the F type approximation can be easily computed.

The quantities Σ_2, Σ_3 and Σ_4 required in these approximations are given by (81), the calculations of Table 9 give some indication of the accuracy to be expected.

6·2. *Special cases*

We consider two important special cases of the statistic, that in which there are only two groups of variates and that in which there is only one variate in each of the k groups.

6·21. *Case $k = 2$*

In this case the expressions for the coefficients in the series simplify considerably. Writing $p_1 = p$, $p_2 = q$, we obtain

$$f = pq, \quad \rho = 1 - \frac{p+q+1}{2\nu}, \quad \beta = \tfrac{1}{2}(p+q+1),$$

$$\alpha_1 = 0, \quad \alpha_2 = \frac{pq}{48}(p^2+q^2-5),$$

$$\alpha_3 = 0, \quad \alpha_4 = \frac{pq}{1920}\{3p^4+3q^4+10p^2q^2-50(p^2+q^2)+159\},$$

$$\alpha_5 = 0, \quad \alpha_6 = \frac{pq}{16{,}128}\{3(p^6+q^6)-105(p^4+q^4)+1{,}113(p^2+q^2) + (21p^2-350+21q^2)p^2q^2-2{,}995\}.$$

Putting these values in (26) and (30) we confirm* the series given to terms in μ^{-4} by Rao (1948) for this case, $k = 2$. Bartlett had already (1938) obtained the χ^2 approximation using the scale factor $\frac{1}{C} = 1 - \frac{p+q+1}{2\nu}$ (which is of course the factor given by the present procedure). Rao introduced this scale factor into the Wald & Brookner series, so as to obtain a χ^2 series with Bartlett's χ^2 approximation as the leading term, equivalent to (30). As we have seen, this choice of factor results in this particular case in α_1 and the α's of odd order being zero, so that the calculation of the series is correspondingly simpler.

* There appears to be a misprint in Rao's paper in the expression which corresponds with α_4, where the constant 159 is wrongly given as 150.

χ^2 (*Bartlett, approximation.* $\frac{M}{C} = \left(1 - \frac{p+q+1}{2\nu}\right) M$ is approximately distributed as χ^2 with $f = pq$ degrees of freedom.

F approximation. We find $A_2 - A_1^2 = (p^2+q^2-5)/12\nu^2$; thus for p and $q \geqslant 2$, $A_2 - A_1^2 > 0$, and the type VI form will be appropriate. M/b will be approximately distributed as F with f_1 and f_2 degrees of freedom, where

$$f_1 = pq, \quad f_2 = \frac{12\nu^2(pq+2)}{p^2+q^2-5}, \quad b = \frac{pq}{1 - \frac{p+q+1}{2\nu} - \frac{f_1}{f_2}}.$$

For p or q equal to 1 and 2, the exact distributions are known and provide simple tests (Wilks, 1932, 1935). For these cases λ and $\sqrt{\lambda}$ respectively are distributed in a type I distribution, and the significance test can be made, either by directly entering Thompson's tables of percentage points of the incomplete B-function, or by inversion of the statistic to its equivalent 'variance ratio' form and using tables of F or of Fisher's z (Bartlett, 1934; Rao, 1948). As has been pointed out by Bartlett (1938) if $p = 1$ and $q = 2$ (or $p = 2$ and $q = 1$) $\frac{M}{C} = \frac{\nu - 2}{\nu} M$ is distributed *exactly* as χ^2, and substituting these values for p and q in the expressions for α_r we find that in this case all these coefficients are zero, so providing a useful check. If p and q were both unity, $A_2 - A_1^2$ would be negative and the type I form be appropriate for the F approximation. Of course we shall not need to use the method here because the criterion $\sqrt{(1-\lambda)}$ is the sample correlation coefficient r and the exact distribution is known. The exact distributions are also known in certain other cases (Wilks, 1935; Wald & Brookner, 1941); the form which these take, however, is rather complicated, but they are useful to check approximations. In Table 10 are shown the 5 % significance points of M for a number of combinations of p and q as given by the χ^2 and F methods of approximation. In the cases chosen, the exact distribution is known, and this has been used to calculate the exact probability associated with each of these points. For comparison, the probability given by the series, using terms up to α_6 in the asymptotic series and, for most values, up to α_8 in the exponential series, is also shown.

We see that, providing ν is sufficiently large, Bartlett's approximation is in good agreement with the exact values, and the F approximation, since it involves very little more labour, provides a worth-while improvement. If ν is not large and one is doubtful whether these approximations will be sufficient, a rough but useful indication is provided by comparing the values obtained by the χ^2 and F approximations (in calculating the F approximation one will have already calculated the quantities needed for the χ^2 approximation). If these two approximations give substantially the same value, it may generally be taken as an indication that the approximation is adequate. If they differ markedly, a more accurate value should be calculated from the series.

6·22. *Case* $p_l = 1$, $l = 1, 2, \ldots, k$

If the k groups each contain only one variate, the hypothesis tested is that each of the variates is independent of all the others. The λ criterion then becomes the determinant of the sample correlation matrix, e.g., if $k = 3$,

$$\lambda = \begin{vmatrix} 1 & r_{12} & r_{13} \\ r_{12} & 1 & r_{23} \\ r_{13} & r_{23} & 1 \end{vmatrix},$$

Table 10. 5 % *significance points for M*

p	q	ν	χ^2 (Bartlett)			F (Box)		
			M	Probability		M	Probability	
				Exact	Series		Exact	Series
1	1	9	4·610	0·0494	0·0494	4·592	0·0499	0·0499
1	5	10	17·032	0·0624	0·0624	17·542	0·0555	0·0555
		20	13·419	0·0518	0·0518	13·504	0·0504	0·0504
1	10	20	26·153	0·0666	0·0666	27·022	0·0562	0·0562
		40	21·538	0·0525	0·0525	21·690	0·0505	0·0505
2	2	9	13·137	0·0515	0·0515	13·200	0·0506	0·0506
2	5	10	30·512	0·0737	0·0737	31·654	0·0614	0·0614
		20	22·884	0·0529	0·0529	23·053	0·0507	0·0507
2	10	20	46·534	0·0753	0·0753	48·164	0·0595	0·0595
		40	37·505	0·0535	0·0535	37·775	0·0507	0·0507
4	4	10	47·811	0·0945	0·0940	49·996	0·0735	0·0731*
		20	33·931	0·0542	0·0542	34·216	0·0512	0·0512

* With $\nu = 10$ and $p = q = 4$, six terms were taken in the asymptotic series and twelve in the exponential series; greater accuracy can be obtained by taking more terms in the asymptotic series.

where r_{ij} is the usual sample product moment correlation coefficient between the ith and jth variates. When $k = 2$ the criterion is simply $1 - r_{12}^2$.

The statistic is useful in supplying an overall test of independence between the k variates. For example, when $k = 5$ there will be ten individual correlation coefficients. Even when the null hypothesis, that all the variates are uncorrelated, is true, we shall expect often to come across individual coefficients which are 'significant'. For such a case it will be appropriate to apply the overall test before testing individual correlations. Again, the expressions for the coefficients simplify and we find, choosing ρ so that $\alpha_1 = 0$,

$$f = \tfrac{1}{2}k(k-1), \quad \rho = 1 - \frac{2k+5}{6\nu}, \quad \beta = \frac{2k+5}{6},$$

$$\alpha_1 = 0, \quad \alpha_2 = \frac{k(k-1)}{288}(2k^2 - 2k - 13),$$

$$\alpha_3 = \frac{k(k-1)}{3{,}240}(k-2)(2k-1)(k+1),$$

$$\alpha_4 = \frac{k(k-1)}{34{,}560}(16k^4 - 32k^3 - 252k^2 + 268k + 1147),$$

$$\alpha_5 = \frac{k(k-1)}{136{,}080}(k-2)(k+1)(2k-1)(8k^2 - 8k - 97),$$

$$\alpha_6 = \frac{k(k-1)}{7{,}838{,}208}(496k^6 - 1{,}488k^5 - 12{,}576k^4 + 27{,}632k^3 + 137{,}490k^2 - 151{,}554k - 562{,}103).$$

For the χ^2 approximation we find, from the argument of §3·1, that we should take $\frac{1}{C} = 1 - \frac{2k+5}{6\nu}$. Thus $\left\{1 - \frac{2k+5}{6\nu}\right\} M$ will be approximately distributed as χ^2 with $\frac{1}{2}k(k-1)$ degrees of freedom.

For the F approximation we have

$$f = \tfrac{1}{2}k(k-1), \quad A_1 = \frac{2k+5}{6\nu}, \quad A_2 = \frac{k^2+3k+2}{6\nu^2}.$$

For $k = 2$ and 3 we use the type I form and for $k \geqslant 4$ the type VI, since $A_2 - A_1^2 = \frac{2k^2 - 2k - 13}{36\nu^2}$ is negative when $k = 2$ or 3 and positive for larger values of k. We then calculate f_1, f_2 and b, required in this approximation, by formulae (77) and (76) respectively.

7. Summary and conclusions

For a particular class of likelihood criteria, whose moments appear as the product of Γ-functions, a general method is described for obtaining probability levels when the null hypothesis is true. A number of statistics whose moments appear in this form are referred to, and a general method developed to obtain:

(*a*) A series which is in close agreement with the exact distribution.

(*b*) An approximate solution, using a single χ^2 distribution, which is sufficiently accurate for moderate or large samples.

(*c*) A rather better approximation, using a single F distribution, giving close agreement even when the samples are rather small.

The method is illustrated for the following two general statistics:

(1) *Tests for constancy of variance and covariance*

(*a*) *Univariate case.* The F approximation is of the same order of accuracy as Hartley's (1940) series solution although it requires very much less calculation, and significance may be judged by consulting tables of the significance points of the variance ratio F alone.

(*b*) *Multivariate case.* The series solution shows remarkably close agreement with the exact distribution when this is known, and is used in other cases to compare approximations. The χ^2 approximation does not correspond with that found by Bishop (1939), but is, in fact, simpler and more accurate.

The series confirms the accuracy of significance points found by fitting a type I curve to the first two moments of l_1. The calculation of the moments involved in this method renders it too laborious for routine use, and Bishop suggested two working approximations; the F approximation developed here is more accurate than these approximations, whilst it involves no more labour and can be used when the sample sizes are unequal.

(2) *Wilks's test for independence of k groups of variates*

The asymptotic series, and χ^2 and F approximations are derived for this case, and the relation of the results with those of Wald & Brookner, Bartlett, and Rao is discussed. The exact distribution is used to assess the accuracy of the proposed methods in a number of cases. The probabilities given by the series are found to be in excellent agreement with the true values, even for fairly small samples. Providing the sample sizes are not too small, the χ^2 and F approximations will be sufficiently accurate, the latter providing the better approximation, and allowing the sample size to be rather smaller than is possible with the χ^2 approximation. When the number of variates in each group is one, we have a test criterion for the hypothesis that k variates are mutually independent, and the same procedure provides the series solution and simple approximations for tests of significance.

In conclusion, I wish gratefully to acknowledge the help and guidance I have received from Dr H. O. Hartley throughout this investigation.

REFERENCES

Barnes, E. W. (1899). *Mess. Math.* p. 64.
Bartlett, M. S. (1934). *Proc. Camb. Phil. Soc.* **30**, 327.
Bartlett, M. S. (1937). *Proc. Roy. Soc.* A, **160**, 268.
Bartlett, M. S. (1938). *Proc. Camb. Phil. Soc.* **34**, 33.
Bartlett, M. S. (1947). *J.R. Statist. Soc. Suppl.* **9**, 176.
Bishop, D. J. (1939). *Biometrika*, **31**, 31.
Bishop, D. J. & Nair, U. S. (1939). *J.R. Statist. Soc. Suppl.* **6**, 89.
Brown, G. W. (1939). *Ann. Math. Statist.* **10**, 119.
George, A. (1945). *Sankhyā*, **1**, 20.
Hartley, H. O. (1940). *Biometrika*, **31**, 249.
Hartley, H. O. & Pearson, E. S. (prefatory note) (1946). *Biometrika*, **33**, 296.
Mauchly, J. W. (1940). *Ann. Math. Statist.* **11**, 204.
Milne-Thomson, L. M. (1933). *The calculus of finite differences.* Macmillan.
Nair, U. S. (1939). *Biometrika*, **30**, 274.
Nayer, P. P. N. (1936). *Statist. Res. Mem.* **1**, 38.
Neyman, J. & Pearson, E. S. (1928). *Biometrika*, **20A**, 175 and 263.
Neyman, J. & Pearson, E. S. (1931). *Bull. int. Acad. Cracovie*, A, p. 460.
Neyman, J. & Pearson, E. S. (1936). *Statist. Res. Mem.* **1**, 1.
Neyman, J. & Pearson, E. S. (1938). *Statist. Res. Mem.* **2**, 25.
Pearson, E. S. & Wilks, S. S. (1933). *Biometrika*, **25**, 353.
Pitman, E. J. G. (1939). *Biometrika*, **31**, 200.
Plackett, R. L. (1946). *J.R. Statist. Soc.* **109**, 457.
Plackett, R. L. (1947). *Biometrika*, **34**, 311.
Rao, C. R. (1948). *Biometrika*, **35**, 71.
Thompson, C. M. (1941). *Biometrika*, **32**, 151.
Thompson, C. M. & Merrington, M. (1946). *Biometrika*, **33**, 296.
Tukey, J. W. & Wilks, S. S. (1946). *Ann. Math. Statist.* **17**, 318.
Votaw, D. F. (1948). *Ann. Math. Statist.* **19**, 447.
Wald, A. & Brookner, R. J. (1941). *Ann. Math. Statist.* **12**, 137.
Welch, B. L. (1935). *Biometrika*, **27**, 145.
Welch, B. L. (1936). *Statist. Res. Mem.* **1**, 52.
Wilks, S. S. (1932). *Biometrika*, **24**, 471.
Wilks, S. S. (1935). *Econometrica*, **3**, 309.
Wilks, S. S. (1946). *Ann. Math. Statist.* **17**, 257.
Wishart, J. (1928). *Biometrika*, **20A**, 32.

NOTE ON APPROXIMATIONS TO THE POWER FUNCTION OF THE '2×2 COMPARATIVE TRIAL'

By G. P. SILLITTO

Research Department, Imperial Chemical Industries, Ltd., Nobel Division

The power function of the 2×2 table arising in what Barnard (1947) has called the '2×2 comparative trial' and Pearson (1947) has designated 'Problem II' has been discussed by P. B. Patnaik (1948) in this journal. In his investigation he made approximations of the type which involve representing binomial or hypergeometric distributions by normal distributions, and in certain cases he examined the adequacy of the approximations numerically, by comparing values of the approximate power function he derived, with values calculated exactly.

There is some interest in comparing Patnaik's approximation with that obtained by using the angular transformation for a binomial variate. If P is the probability that an individual will possess a given character and $f = r/n$ is the proportion or relative frequency of individuals with this character observed in a random sample of n, then it is known that

$$x = \arcsin \sqrt{f}, \tag{1}$$

where the angle is measured in radians, is distributed approximately normally about a mean of $\arcsin \sqrt{P}$ with a standard deviation of $\frac{1}{2}n^{-\frac{1}{2}}$. The problem of comparing two observed values of f, thus becomes equivalent to that of comparing variables from normal populations with known standard deviations.

The transformation (1) has of course been known for a long time. It has recently been discussed by Eisenhart (1947), who gives a bibliography, and Paulson & Wallis (1947) have indicated its application in the planning and analysing of experiments for comparing two percentages. Bartlett (1937) has given a table of the transformation, the angle being in radians, while Fisher & Yates (1938) have tabulated it, showing the angle in degrees. Since the reason for using a transformation is essentially one of convenience and since the use of the radian as the unit leads to a slightly simpler expression for the variance, there are advantages in employing the latter unit. For the case of two independent samples, if the observed relative frequencies f_1 and f_2 are based respectively on m and n observations, then the difference between their respective transformations x_1 and x_2 will be approximately normally distributed with known standard error,

$$\sigma = \frac{1}{2}\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)}. \tag{2}$$

Thus $(x_2 - x_1)/\sigma$ will vary about a mean

$$\mu = (\arcsin \sqrt{P_2} - \arcsin \sqrt{P_1})/\sigma, \tag{3}$$

with unit standard deviation. If the null hypothesis is true, $P_1 = P_2$ and $\mu = 0$. Defining u_α as follows,

$$\alpha = \int_{u_\alpha}^{\infty} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}t^2} dt, \tag{4}$$

the chance of establishing significance at the 100α % level, when $P_1 \neq P_2$, i.e. the power of the test, will be approximated as follows:

(*a*) For one-sided test $$\text{Power} = \int_{u_\alpha - \mu}^{\infty} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}t^2}\, dt. \tag{5}$$

(*b*) For two-sided test

$$\text{Power} = \int_{-\infty}^{-u_{\frac{1}{2}\alpha}-\mu} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}t^2}\, dt + \int_{u_{\frac{1}{2}\alpha}-\mu}^{\infty} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}t^2}\, dt. \tag{6}$$

The position is illustrated in Fig. 1. For the one-sided test the power is the area under the curve centred at μ lying to the right of the ordinate at u_α. For the two-sided test it is the sum of the areas under this curve lying to the left of the ordinate at $-u_{\frac{1}{2}\alpha}$ and to the right of the ordinate at $u_{\frac{1}{2}\alpha}$.

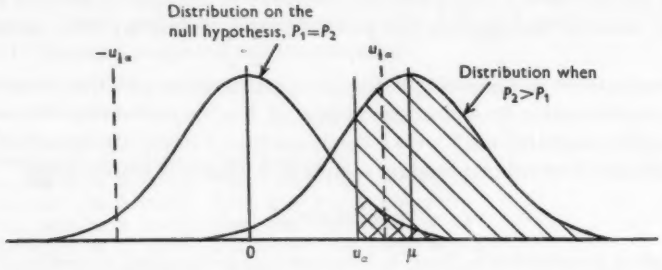

Fig. 1. Sampling distributions of $(x_2-x_1)/\sigma = \{\text{arc sin } \sqrt{f_2} - \text{arc sin } \sqrt{f_1}\} \Big/ \left\{\frac{1}{2}\sqrt{\left(\frac{1}{m}+\frac{1}{n}\right)}\right\}$.

In order to obtain an indication of the usefulness of the approximation to the power of the 2×2 comparative trial which is afforded by using the transformation (1) and the normal distribution theory just outlined, values for the power of the test have been calculated for the cases for which Patnaik evaluated the true power of the 2×2 table. The results are given in Tables 1 (*a*), (*b*), 2 (*a*) and (*b*) below. The first two tables correspond to Patnaik's Tables 2 (*a*) and (*b*), and the bracketed figures are the values for the power as calculated from the approximate theory given above, while the other figures are the values for the true power, given by Patnaik. Similarly, in Tables 2 (*a*) and (*b*), the exact values are quoted from Tables 3 (*a*) and (*b*) in Patnaik's paper, and the corresponding approximate values from the present method are shown for comparison.

It is evident that the approximation is quite good, over the range covered by the present tables. It appears, in fact, to be slightly better on the whole than Patnaik's, a detailed comparison of the figures with those in his paper giving the results shown in Table 3. Omitting the $P_1 = P_2$ diagonals of Tables 1 (*a*) and (*b*), there are 172 cases in which numerical comparison is possible.

It may be mentioned also that in the region where the true power is 0·9 or greater, which is probably the most important in practice, Patnaik's approximations generally tend to overestimate the power, whilst the present method tends to under-estimate it. Some indication of the practical importance of the difference between the approximations to the power function can be obtained in cases where the true power is known, by using the approximations to provide estimates of n, the sample size required in order to afford a test of assigned power,

Table 1 (*a*). *Power function for* $m = 18$, $n = 12$.
Significance level $\alpha = 0{\cdot}10$ *in the two-sided test*

Approximate values derived through the angular transformation are shown in parentheses.

P_2 \ P_1	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9
0·9	1·000 (1·000)	0·997 (0·995)	0·975 (0·974)	0·907 (0·917)	0·773 (0·800)	0·574 (0·619)	0·356 (0·398)	0·170 (0·197)	0·088 (0·100)
0·8	0·997 (0·995)	0·972 (0·965)	0·882 (0·882)	0·732 (0·733)	0·494 (0·533)	0·289 (0·326)	0·144 (0·165)	0·094 (0·100)	0·202
0·7	0·977 (0·974)	0·880 (0·882)	0·722 (0·714)	0·487 (0·501)	0·272 (0·298)	0·136 (0·153)	0·092 (0·100)	0·159	0·393
0·6	0·918 (0·917)	0·742 (0·733)	0·515 (0·501)	0·287 (0·290)	0·142 (0·159)	0·099 (0·100)	0·155	0·320	0·610
0·5	0·796 (0·800)	0·534 (0·533)	0·309 (0·298)	0·156 (0·159)	0·100 (0·100)	0·156	0·309	0·534	0·796
0·4	0·610 (0·619)	0·320 (0·326)	0·155 (0·153)	0·099 (0·100)	0·142	0·287	0·515	0·742	0·918
0·3	0·393 (0·398)	0·159 (0·165)	0·092 (0·100)	0·136	0·272	0·487	0·722	0·880	0·977
0·2	0·202 (0·197)	0·094 (0·100)	0·144	0·289	0·494	0·732	0·882	0·972	0·997
0·1	0·088 (0·100)	0·170	0·356	0·574	0·773	0·907	0·975	0·997	1·000

Table 1 (*b*). *Power function for* $m = 18$, $n = 12$.
Significance level $\alpha = 0{\cdot}02$ *in the two-sided test*

Approximate values derived through the angular transformation are shown in parentheses.

P_2 \ P_1	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9
0·9	0·998 (0·996)	0·976 (0·971)	0·902 (0·897)	0·752 (0·759)	0·532 (0·564)	0·290 (0·353)	0·103 (0·173)	0·017 (0·060)	0·006 (0·020)
0·8	0·961 (0·971)	0·882 (0·870)	0·694 (0·693)	0·490 (0·476)	0·265 (0·275)	0·115 (0·127)	0·035 (0·046)	0·016 (0·020)	0·055
0·7	0·909 (0·897)	0·714 (0·693)	0·460 (0·453)	0·248 (0·248)	0·111 (0·111)	0·041 (0·041)	0·021 (0·020)	0·050	0·177
0·6	0·767 (0·759)	0·500 (0·476)	0·262 (0·248)	0·111 (0·107)	0·041 (0·039)	0·023 (0·020)	0·048	0·138	0·361
0·5	0·625 (0·564)	0·293 (0·275)	0·124 (0·111)	0·045 (0·039)	0·022 (0·020)	0·045	0·124	0·293	0·625
0·4	0·361 (0·353)	0·138 (0·127)	0·048 (0·041)	0·023 (0·020)	0·041	0·111	0·262	0·500	0·767
0·3	0·177 (0·173)	0·050 (0·046)	0·021 (0·020)	0·041	0·111	0·248	0·460	0·714	0·909
0·2	0·055 (0·060)	0·016 (0·020)	0·035	0·115	0·265	0·490	0·694	0·882	0·961
0·1	0·006 (0·020)	0·017	0·103	0·290	0·532	0·752	0·902	0·976	0·998

Table 2 (*a*). *Some values of the power function for the two-sided test for* $m = n = 15$

P_1	P_2	$\alpha = 0{\cdot}10$		$\alpha = 0{\cdot}02$	
		Exact value	Approximate value from angular transformation	Exact value	Approximate value from angular transformation
0·3	0·4	0·141	0·156	0·034	0·042
0·6	0·8	0·306	0·334	0·112	0·133
0·1	0·3	0·389	0·409	0·149	0·180
0·2	0·7	0·896	0·893	0·680	0·713
0·05	0·5	0·916	0·922	0·736	0·770
0·1	0·6	0·919	0·926	0·739	0·778
0·2	0·8	0·974	0·970	0·872	0·885
0·1	0·7	0·980	0·978	0·894	0·910

Table 2 (*b*). *Some values of the power function for the two-sided test for* $m = n = 30$

P_1	P_2	$\alpha = 0{\cdot}10$		$\alpha = 0{\cdot}02$	
		Exact value	Approximate value from angular transformation	Exact value	Approximate value from angular transformation
0·05	0·3	0·884	0·864	0·631	0·662
0·1	0·4	0·885	0·878	0·691	0·686
0·3	0·7	0·937	0·939	0·807	0·805
0·2	0·6	0·945	0·948	0·839	0·828
0·1	0·5	0·977	0·974	0·902	0·897
0·2	0·7	0·993	0·993	0·965	0·961

Table 3. *Comparison of approximations to the power function of the* 2×2 *comparative trial*

Table no.	No. of cases in which Patnaik's first approximation is nearer to the exact value	No. of indecisive cases (equally discrepant, or indeterminate owing to rounding-up)	No. of cases in which the angular transformation is nearer to the exact value
1 (*a*)	6	28	38
1 (*b*)	6	36	30
2 (*a*)	0 (1)	3 (3)	13 (12)
2 (*b*)	0 (0)	0 (1)	12 (11)
Totals	12	67	93

Bracketed figures in the body of this table refer to Patnaik's second approximation.

β, when a particular significance level, α, has been chosen. In his Table 7 Patnaik has estimated n for selected cases of this kind using the two-sided test and $\alpha = 0{\cdot}10$ and $0{\cdot}02$. Table 4 contains his results together with those obtained by using the angular transformation approximation. In order to obtain a simple expression for n, using this approximation, the left-hand term in expression (6) may be neglected, since it makes an appreciable contribution to the power only when the power is rather low, unless α is larger than is usual in significance testing. Making the relation between β and u_β, as for α and u_α defined in equation (4), it follows that for the two-sided test

$$\mu = u_{\frac{1}{2}\alpha} - u_\beta. \tag{7}$$

Hence, using equations (2) and (3), it is seen that

$$\frac{1}{m} + \frac{1}{n} = 4(\arcsin\sqrt{P_2} - \arcsin\sqrt{P_1})^2/(u_{\frac{1}{2}\alpha} - u_\beta)^2. \tag{8}$$

In the particular case when $m = n$, therefore, n is given by

$$n = \tfrac{1}{2}(u_{\frac{1}{2}\alpha} - u_\beta)^2/(\arcsin\sqrt{P_2} - \arcsin\sqrt{P_1})^2. \tag{9}$$

This formula has been given by Paulson & Wallis (1947) with reference to the single-sided test, and they have provided a nomogram for determining sample sizes in the one-sided case, which can, of course, be used for a two-sided experiment if their α is taken as one-half the risk of the error of the first kind which can be tolerated. Using (9) on the cases of Patnaik's Table 7, the values in Table 4 are obtained.

Table 4. *Estimation of the sample size from Patnaik's first approximation and from the angular transformation approximation*

True sample size $= n$	P_1	P_2	True power*		Patnaik's estimate of n		Estimate of n from equation (9)	
			For $\alpha = 0{\cdot}10$	For $\alpha = 0{\cdot}02$	For $\alpha = 0{\cdot}10$	For $\alpha = 0{\cdot}02$	For $\alpha = 0{\cdot}10$	For $\alpha = 0{\cdot}02$
15	0·05	0·5	0·916	0·736	11	11	15	14
15	0·1	0·6	0·919	0·739	13	12	15	14
15	0·1	0·7	0·980	0·894	13	12	16	15
30	0·05	0·3	0·884	0·631	26	23	33	29
30	0·1	0·5	0·977	0·902	27	26	31	31
30	0·2	0·7	0·993	0·965	28	28	31	31

* The power of the two-sided test.

The angular transformation (1) is, of course, a particular case of a type of transformation recently considered by Anscombe (1948), and it is not difficult to show, using his methods, that for finite numbers of observations and $P \neq 0{\cdot}5$, the expectation of x is not exactly $\arcsin\sqrt{P}$, but differs from it by a function of n and P, which increases as n decreases and as $|P - 0{\cdot}5|$ increases. The higher moments of x exhibit rather similar departures from those of a normal variate. It is to be expected, therefore, that the angular transformation approxi-

mation will be unlikely to remain satisfactory for smaller n and extreme frequencies, but on the present evidence it appears to be satisfactory for values of n and P which are of interest in many practical problems.

Grateful acknowledgement is made to Prof. E. S. Pearson for his interest and for helpful suggestions as to the framing of this note.

REFERENCES

ANSCOMBE, F. J. (1948). The transformation of Poisson, binomial, and negative-binomial data. *Biometrika*, **35**, 246.

BARNARD, G. A. (1947). Significance tests for 2 × 2 tables. *Biometrika*, **34**, 123.

BARTLETT, M. S. (1937). Sub-sampling for attributes. *J.R. Statist. Soc. Suppl.* **4**, 131.

EISENHART, C. (1947). Inverse Sine Transformation of Proportions. Chapter 16 of *Selected Techniques of Statistical Analysis* by the Statistical Research Group, Columbia University; edited by Eisenhart, Hastay and Wallis. New York: McGraw-Hill Book Co. Inc.

FISHER, R. A. & YATES, F. (1938). *Statistical Tables for Biological, Agricultural and Medical Research*, 3rd ed. 1948, p. 56. Edinburgh: Oliver and Boyd.

PATNAIK, P. B. (1948). The power function of the test for the difference between two proportions in a 2 × 2 table. *Biometrika*, **35**, 157.

PAULSON, E. & WALLIS, W. A. (1947). Planning and Analyzing Experiments for comparing Two Percentages. Chapter 7 of *Selected Techniques of Statistical Analysis* by the Statistical Research Group, Columbia University; edited by Eisenhart, Hastay and Wallis. New York: McGraw-Hill Book Co. Inc.

PEARSON, E. S. (1947). The choice of statistical tests illustrated on the interpretation of data classed in a 2 × 2 table. *Biometrika*, **34**, 139.

THE DISTRIBUTION OF 'STUDENT'S' t IN RANDOM SAMPLES OF ANY SIZE DRAWN FROM NON-NORMAL UNIVERSES

By A. K. GAYEN, *St Catharine's College, University of Cambridge*

1. Introduction

The effect of universal non-normality on 'Student's' t has been studied so far either by way of particular numerical examples or by approximations to its sampling distribution based on the large sample assumption. E. S. Pearson (1928, 1929) has shown in his experimental investigation that the effect of universal 'excess' and of 'skewness' on 'Student's' ratio z (which is related to t by $t = z\sqrt{(n-1)}$) may be considerable. He has also furnished a small table, based on experimental results, showing some actual values of the probability integral for the ratio in samples of 2, 5, 10 and 20 from a few universes with specified values of 'skewness' and 'excess'. M. S. Bartlett (1935) obtained some theoretical results for approximately representing the distribution of t, taking into account the universal 'skewness' and 'excess'. The approximation is based on the assumption of large samples and is not very satisfactory, for in that approach, as he himself has also pointed out, the term 'approximation' is perhaps misleading. Still from the form of the expression obtained, it has been observed that the effect of the 'skewness' λ_3 in the original distribution is the more serious and that of the 'excess' λ_4 is small. This, of course, confirms Pearson's experimental results, but the resulting expression cannot clearly furnish a quantitative measure of the corrections to be applied to normal theory probabilities of t, when the sample size is small. R. C. Geary (1936) obtained an expression for the distribution of t in samples of any size drawn from a slightly asymmetrical universe. It consists of two components, one being the 'normal theory' t and the other a term in λ_3, which has been called the 'corrective function' due to it. He has also given a table of corrective tail-area probabilities for some representative values of n. In a recent communication (1947) he has given from a different approach an approximate formula for the frequency of t, correct to n^{-2}, n being the size of sample. From this and analogous results on non-normal Fisher's z, he has shown, by a few illustrative examples, that the inferences drawn from the standard tables may be seriously in error even in some cases where the parent is not considerably non-normal. He has accordingly pointed out the primary importance of testing for normality from the available samples and has suggested that when universal normality cannot be assumed, the standard tables should be corrected by using the sample estimates of λ_3 and λ_4, etc., in conjunction with the theoretical results deduced therein. But so far as the frequency of t is concerned, satisfactory measures of probabilities for small samples cannot be obtained from his results, since they are mainly based on large sample assumptions. Also the first few terms of his result, as far as n^{-1}, agree with those of Bartlett, which, as has been pointed out, are not very satisfactory.

The purpose of the present investigation is to obtain the form of the corrective terms due to population cumulants λ_3, λ_4, etc., in the frequency function of t for any size of sample. For this the parent population will be specified by the Edgeworth series. It has been proved rigorously by Cramér (1928) that this series gives a real asymptotic expansion of any universe $f(x)$, in powers of $\nu_0^{-\frac{1}{2}}$, where ν_0 is the number of sources of 'elementary errors'. Inclusion of

terms of order $\nu_0^0, \nu_0^{-\frac{1}{2}}, \nu_0^{-1}, \ldots, \nu_0^{-\frac{1}{2}r}$ gives rise to the first, second, third, ..., $(r+1)$th approximations to the law of error. We shall consider mainly the third approximation, including terms in λ_3, λ_4 and λ_3^2, but shall also deal with the fourth approximation which takes in λ_5, $\lambda_3\lambda_4$ and λ_3^3.

Probabilities of values of t, obtained from the derived formula, for the parent specified by the third approximation, are in satisfactory agreement with the experimental determinations made so far, even in cases where the sampled populations are represented by the exponential curve or a very skew Type III curve of Pearson. Accordingly, it appears that the derived expressions for 'Student's' t may perhaps provide quite satisfactory estimates of the probabilities, for a fairly wide class of non-normal universes, especially when the sample size is not too small.

The values of the cumulants of the population have been assumed to be known, and the question of their estimation has not been considered.

2. Joint distribution of the sum and the sum of squares of deviations of n sample observations

We shall first of all obtain the distribution of 'Student's' t, with particular reference to the parent population being specified by the third approximation to the law of error. This will give us the corrective terms due to λ_3, λ_4 and λ_3^2. It will be shown afterwards that similar corrective terms due to λ_5, $\lambda_3\lambda_4$ and λ_3^3 can be obtained by the same methods when the parent population is represented by the Edgeworth series up to the corresponding terms.

Let us then consider the parent population to be specified by the third approximation, correct to ν_0^{-1},

$$f(x) = \phi(x) - \frac{\lambda_3}{3!}\phi^{(3)}(x) + \frac{\lambda_4}{4!}\phi^{(4)}(x) + \frac{10\lambda_3^2}{6!}\phi^{(6)}(x), \tag{2·1}$$

where λ_3 and λ_4 are the measures of universal skewness and excess,

$$\phi(x) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}x^2}, \tag{2·2}$$

the normal function in standardized form, and

$$\phi^{(\nu)}(x) = \left(\frac{d}{dx}\right)^\nu \phi(x) = (-1)^\nu H_\nu(x)\phi(x), \tag{2·3}$$

where $H_\nu(x)$ is the well-known Hermitian polynomial of degree ν.

Let $x_1, x_2, \ldots, x_n$ be n independent sample observations drawn at random from the universe (2·1), and let us adopt the notations

$$S_1 = \sum_1^n x_\nu, \quad S_2' = \sum_1^n x_\nu^2 \tag{2·4}$$

and

$$S_2 = \sum_1^n (x_\nu - \bar{x})^2 = S_2' - \frac{S_1^2}{n} = (n-1)s^2, \tag{2·5}$$

where s^2 is the estimated variance.

By the use of Bartlett's method (1935) the joint frequency density of (S_1, S_2) (which is also the joint frequency density of (S_1, S_2'), since the Jacobian of transformation is unity) may be written in the form

$$g_1(S_1, S_2) = W(n-1) - \frac{n\lambda_3}{3!}\{D_1^3 W(n+5) - 6D_1 D_2 W(n+3)\}$$

$$+ \frac{n\lambda_4}{4!}\{D_1^4 W(n+7) - 12D_1^2 D_2 W(n+5) + 12D_2^2 W(n+3)\}$$

$$+ \frac{n\lambda_3^2}{72}\{nD_1^6 W(n+11) - 6(2n+3)\, D_1^4 D_2 W(n+9) + 36(n+4)\, D_1^2 D_2^2 W(n+7) - 120 D_2^3 W(n+5)\}, \tag{2·6}$$

when terms in λ_i other than in λ_3, λ_4 and λ_3^2 are neglected. In (2·6)

$$W(n_0') = \left\{\frac{e^{-S_1^2/2n}}{\sqrt{(2\pi n)}}\right\}\left\{\frac{S_2^{\frac{1}{2}(n_0'-2)} e^{-\frac{1}{2}S_2}}{2^{\frac{1}{2}n_0'}\,\Gamma(\frac{1}{2}n_0')}\right\}$$

$$= \frac{e^{-\frac{1}{2}S_1^2}\,(S_2' - S_1^2/n)^{\frac{1}{2}(n_0'-2)}}{\sqrt{(2\pi n)} \times 2^{\frac{1}{2}n_0'}\,\Gamma(\frac{1}{2}n_0')}, \tag{2·7}$$

and

$$D_1 \equiv \frac{\partial}{\partial S_1}, \quad D_2 \equiv \frac{\partial}{\partial S_2'}. \tag{2·8}$$

We now proceed to evaluate the derivatives involved in the expression (2·6). It is possible to transform the product operators of the form $D_1^{\nu_1} D_2^{\nu_2}$ (where $\nu_1, \nu_2 = 0, 1, 2, \ldots$) into functions of the single operator D_1.

For any typical function $W(n_0')$, defined in (2·7), where n_0' stands for $(n-1)$, $(n+1)$, $(n+3)$, ..., it can be shown that

$$D_1^{\nu_1} D_2^{\nu_2} W(n_0') = \frac{(-1)^{\nu_2}}{2^{\nu_2}}\left[D_1^{\nu_1} W(n_0') - \binom{\nu_2}{1} D_1^{\nu_1} W(n_0'-2) + \binom{\nu_2}{2} D_1^{\nu_1} W(n_0'-4) - \ldots\right.$$
$$\left. + (-1)^r \binom{\nu_2}{r} D_1^{\nu_1} W(n_0'-2r) + \ldots + (-1)^{\nu_2} D_1^{\nu_1} W(n_0'-2\nu_2)\right]. \tag{2·9}$$

By differentiating $W(n_0')$ successively with respect to S_1 we get

$$D_1 W(n_0') = (-1)\, W(n_0'-2)\left(\frac{S_1}{n}\right),$$

$$D_1^2 W(n_0') = (-1)^2\, W(n_0'-4)\left\{\left(\frac{S_1}{n}\right)^2 - \frac{1}{(n_0'-4)}\left(\frac{S_2}{n}\right)\right\},$$

$$D_1^3 W(n_0') = (-1)^3\, W(n_0'-6)\left\{\left(\frac{S_1}{n}\right)^3 - \frac{3}{(n_0'-6)}\left(\frac{S_1}{n}\right)\left(\frac{S_2}{n}\right)\right\},$$

$$D_1^4 W(n_0') = (-1)^4\, W(n_0'-8)\left\{\left(\frac{S_1}{n}\right)^4 - \frac{6}{(n_0'-8)}\left(\frac{S_1}{n}\right)^2\left(\frac{S_2}{n}\right) + \frac{3}{(n_0'-6)\,(n_0'-8)}\left(\frac{S_2}{n}\right)^2\right\},$$

and so on.

It will be observed that the numerical coefficients are those of the Hermitian polynomials

$$H_\nu(x) = x^\nu - \frac{\nu(\nu-1)}{2.1!}\, x^{\nu-2} + \frac{\nu(\nu-1)\,(\nu-2)\,(\nu-3)}{2^2.2!}\, x^{\nu-4} - \ldots.$$

So we can write, in general,

$$D_1^\nu W(n_0') = (-1)^\nu W(n_0'-2\nu)\left\{\left(\frac{S_1}{n}\right)^\nu - \frac{\nu(\nu-1)}{2.1!(n_0'-2\nu)}\left(\frac{S_1}{n}\right)^{\nu-2}\left(\frac{S_2}{n}\right) + \frac{\nu(\nu-1)(\nu-2)(\nu-3)}{2^2.2!(n_0'-2\nu+2)(n_0'-2\nu)}\left(\frac{S_1}{n}\right)^{\nu-4}\left(\frac{S_2}{n}\right)^2 - \ldots\right\}. \quad (2{\cdot}10)$$

For the present case as we have specified the population by the third approximation to the law of error we require the values of D_1^ν, for $\nu = 0, 1, 2, \ldots$ up to 6.

Now using results (2·9) and (2·10) in (2·6), we obtain, after some simplification, the following expression for the joint frequency function of S_1 and S_2:

$$g_1(S_1, S_2) = W(n-1)\left[1 + \frac{n\lambda_3}{3!}\left\{H_3\left(\frac{S_1}{n}\right) + 3\left(\frac{S_2}{n}\right)H_1\left(\frac{S_1}{n}\right)\right\} + \frac{n\lambda_4}{4!}\left\{H_4\left(\frac{S_1}{n}\right) + 6\left(\frac{S_2}{n}\right)H_2\left(\frac{S_1}{n}\right) + 3\frac{(n-1)}{(n+1)}\left(\frac{S_2}{n}\right)^2\right\} + \frac{n\lambda_3^2}{72}\left\{\left[n\left(\frac{S_1}{n}\right)^6 - 3(2n+3)\left(\frac{S_1}{n}\right)^4 + 9(n+4)\left(\frac{S_1}{n}\right)^2 - 15\right] + 6\left(\frac{S_2}{n}\right)\left[n\left(\frac{S_1}{n}\right)^4 - 3(n+3)\left(\frac{S_1}{n}\right)^2 + 6\right] + \frac{9}{(n+1)}\left(\frac{S_2}{n}\right)^2 \times \left[n(n+1)\left(\frac{S_1}{n}\right)^2 - 3(n-1)\right] + 6\frac{n(n-2)}{(n+3)(n+1)}\left(\frac{S_2}{n}\right)^3\right\}\right], \quad (2{\cdot}11)$$

where $H_\nu(S_1/n)$ is the Hermitian polynomial of degree ν in (S_1/n), and $W(n-1)$ is given by (2·7). As a check it was found that, to the required approximation, $\int_0^\infty g_1(S_1, S_2)\,dS_2$ gave the standard formula for the distribution of the sum of n sample observations.

3. The frequency function of estimated variance

Integrating (2·11) for S_1 between the limits $-\infty$ and ∞, we find, for the frequency of S_2, after some simplification,

$$g_1(\cdot, S_2) = V(S_2, (n-1))\left[1 + \frac{\lambda_4}{8n}(n-1)^2\left\{\frac{S_2^2}{(n+1)(n-1)} - 2\frac{S_2}{(n-1)} + 1\right\} + \frac{\lambda_3^2}{12n}(n-1)(n-2)\left\{\frac{S_2^3}{(n+3)(n+1)(n-1)} - \frac{3S_2^2}{(n+1)(n-1)} + \frac{3S_2}{(n-1)} - 1\right\}\right], \quad (3{\cdot}1)$$

where $V(S_2, (n-1))$ is the frequency density of normal S_2. For the distribution of $s^2 = S^2/(n-1)$, this becomes

$$g_1(\cdot, (n-1)s^2)(n-1)\,d(s^2). \quad (3{\cdot}2)$$

Note that if we introduce the operator $D_{02} \equiv \dfrac{\partial}{\partial[S_2/(n-1)]}$, the frequency function (3·1) can be put in the form

$$V(n-1) + \frac{\lambda_4}{2n}D_{02}^2 V(n+3) - \frac{2}{3n}\lambda_3^2\frac{(n-2)}{(n-1)^2}D_{02}^3 V(n+5), \quad (3{\cdot}3)$$

where S_2 has been omitted from the V's.

4. THE DISTRIBUTION OF 'STUDENT'S' t

Put $S_1 = \left(\frac{n}{n-1}S_2\right)^{\frac{1}{2}} t$ in (2·11) and integrate for S_2 from 0 to ∞, when we find that a typical term $W(n-1)S_1^{r_1}S_2^{r_2}\,dS_1 dS_2$ becomes

$$\frac{n^{\frac{1}{2}r_1}.2^{\frac{1}{2}(r_1+2r_2)}\,\Gamma[\frac{1}{2}(n+r_1+2r_2)]}{\sqrt{\pi}\,(n-1)^{\frac{1}{2}(r_1+1)}\,\Gamma[\frac{1}{2}(n-1)]}\,\frac{t^{r_1}\,dt}{[1+t^2/(n-1)]^{\frac{1}{2}(n+r_1+2r_2)}}. \tag{4·1}$$

It will be noticed that (4·1) gives the *normal theory* t for $r_1 = r_2 = 0$. We shall denote the distribution of t by $p(t)$. Thus we get by using (4·1) and after simplification

$$\begin{aligned} p(t) = {} & \frac{\Gamma(\frac{1}{2}n)}{\sqrt{\pi[(n-1)]}\,\Gamma[\frac{1}{2}(n-1)]}\,\frac{1}{[1+t^2/(n-1)]^{\frac{1}{2}n}} + \lambda_3\,\frac{\{3(n-1)t-(2n-1)t^3\}}{6(n-1)\sqrt{(2n\pi)}\,[1+t^2/(n-1)]^{\frac{1}{2}(n+3)}} \\ & - \lambda_4\,\frac{\Gamma[\frac{1}{2}(n+2)]}{24n\sqrt{[\pi(n-1)]}\,\Gamma[\frac{1}{2}(n+3)]}\,\frac{\{3(n-1)-6(n+1)t^2+(n+1)t^4\}}{[1+t^2/(n-1)]^{\frac{1}{2}(n+4)}} \\ & + \lambda_3^2\,\frac{\Gamma[\frac{1}{2}(n+2)]\left\{\begin{array}{l}3(n-1)^2(2n+11)-9(n-1)(n+3)(2n-1)t^2 \\ \quad -3(n+1)(n+3)(2n+13)t^4+(n+1)(n+3)(2n+5)t^6\end{array}\right\}}{144n(n-1)^3\sqrt{\pi}\,\Gamma[\frac{1}{2}(n+5)]\,[1+t^2/(n-1)]^{\frac{1}{2}(n+6)}} \\ = {} & p_0(t)+\lambda_3 p_{\lambda_3}(t)-\lambda_4 p_{\lambda_4}(t)+\lambda_3^2 p_{\lambda_3^2}(t). \end{aligned} \tag{4·2}$$

Obviously $p_0(t)$ is the normal theory t for a sample of size n, and $p_{\lambda_3}(t)$, $p_{\lambda_4}(t)$ and $p_{\lambda_3^2}(t)$ are the corrective terms due to universal λ_3, λ_4 and λ_3^2.

As a check on the above expression for t, moments may be calculated directly, about $t = 0$ and to n^{-2}; they are found to be

$$\left.\begin{aligned} \mu_1'(t) &= -\frac{\lambda_3}{2n^{\frac{1}{2}}}\left(1+\frac{3}{4n}+\ldots\right), \\ \mu_2'(t) &= 1+\frac{2}{n}(1+\lambda_3^2)+\frac{2}{n^2}(3-\lambda_4)+\ldots, \\ \mu_3'(t) &= -\frac{7\lambda_3}{2n^{\frac{1}{2}}}\left(1+\frac{15}{4n}+\ldots\right), \\ \mu_4'(t) &= 3+\frac{2}{n}(9-\lambda_4+14\lambda_3^2)+\frac{1}{n^2}(102-30\lambda_4+120\lambda_3^2)+\ldots. \end{aligned}\right\} \tag{4·3}$$

These are in agreement with the asymptotic formulae for the moments of t for any universe as obtained by Geary (1936), up to first power of λ_4 and second power of λ_3. All these corrective terms when integrated between the limits $-\infty$ and ∞ contribute nothing towards the integral, as may be easily deduced from the fact that $\int_{-\infty}^{\infty} p(t)\,dt = 1$.

5. THE DISTRIBUTION OF 'STUDENT'S' t FOR THE PARENT SPECIFIED BY THE FOURTH APPROXIMATION TO THE LAW OF ERROR

Let us specify the parent universe by the fourth approximation to the law of error, correct up to $\nu_0^{-\frac{1}{2}}$, given by

$$f(x) = \phi(x)-\frac{\lambda_3}{3!}\phi^{(3)}(x)+\frac{\lambda_4}{4!}\phi^{(4)}(x)+\frac{10\lambda_3^2}{6!}\phi^{(6)}(x)-\frac{\lambda_5}{5!}\phi^{(5)}(x)-\frac{35\lambda_3\lambda_4}{7!}\phi^{(7)}(x)-\frac{280\lambda_3^3}{9!}\phi^{(9)}(x). \tag{5·1}$$

It is of interest to know the distribution of s^2 and 'Student's' t in random samples from this universe. We can arrive at these results by following a similar procedure, and the expressions corresponding to (2·6), (2·11) and (4·2) will have additional terms involving λ_5, $\lambda_3\lambda_4$ and λ_3^3. We shall omit these intermediate steps and give below only the extra terms in the joint distribution of S_1 and S_2 and in the distribution of t. Thus corresponding to (2·11), the joint distribution $g_2(S_1, S_2)$, of S_1 and S_2 will be given by

$$
\begin{aligned}
g_2(S_1,S_2) = g_1(S_1,S_2) + W(n-1)\Big[&\frac{n\lambda_5}{5!}\Big\{H_5\Big(\frac{S_1}{n}\Big) + 10\Big(\frac{S_2}{n}\Big)H_3\Big(\frac{S_1}{n}\Big) + 15\frac{(n-1)}{(n+1)}\Big(\frac{S_2}{n}\Big)^2 H_1\Big(\frac{S_1}{n}\Big)\Big\}\\
&+\frac{n\lambda_3\lambda_4}{144}\Big\{\Big[n\Big(\frac{S_1}{n}\Big)^7 + 3(3n+4)\Big(\frac{S_1}{n}\Big)^5 + 21(n+4)\Big(\frac{S_1}{n}\Big)^3 + 3(3n+32)\Big(\frac{S_1}{n}\Big)\Big]\\
&-\frac{1}{(n-1)}\Big(\frac{S_2}{n}\Big)\Big[3n(3n+11)\Big(\frac{S_1}{n}\Big)^5 + 6(7n^2+43n+20)\Big(\frac{S_1}{n}\Big)^3 + 9(3n^2+39n+28)\Big(\frac{S_1}{n}\Big)\Big]\\
&+\frac{1}{(n+1)(n-1)}\Big(\frac{S_2}{n}\Big)^2\Big[3n(7n^2+58n+75)\Big(\frac{S_1}{n}\Big)^3 + 9(3n^3+46n^2+71n+20)\Big(\frac{S_1}{n}\Big)\Big]\\
&-\frac{3n}{(n+3)(n+1)(n-1)}\Big(\frac{S_2}{n}\Big)^3\Big[(3n^3+53n^2+129n+95)\Big(\frac{S_1}{n}\Big)\Big]\Big\}\\
&+\frac{n\lambda_3^3}{1296}\Big\{\Big[n^2\Big(\frac{S_1}{n}\Big)^9 + 9n(n+3)\Big(\frac{S_1}{n}\Big)^7 + 27(n+1)(n+6)\Big(\frac{S_1}{n}\Big)^5\\
&\qquad + 9(n+3)(3n+32)\Big(\frac{S_1}{n}\Big)^3 + 135(n+6)\Big(\frac{S_1}{n}\Big)\Big]\\
&-\frac{9}{(n-1)}\Big(\frac{S_2}{n}\Big)\Big[n^2(n+7)\Big(\frac{S_1}{n}\Big)^7 + 3n(2n^2+21n+33)\Big(\frac{S_1}{n}\Big)^5\\
&\qquad + 3(3n^3+51n^2+166n+60)\Big(\frac{S_1}{n}\Big)^3 + 3(23n^2+161n+96)\Big(\frac{S_1}{n}\Big)\Big]\\
&+\frac{27}{(n+1)(n-1)}\Big(\frac{S_2}{n}\Big)^2\Big[n^2(n^2+14n+41)\Big(\frac{S_1}{n}\Big)^5 + n(3n^3+61n^2+271n+225)\Big(\frac{S_1}{n}\Big)^3\\
&\qquad + 3(13n^3+106n^2+131n+30)\Big(\frac{S_1}{n}\Big)\Big]\\
&-\frac{9n}{(n+3)(n+1)(n-1)}\Big(\frac{S_2}{n}\Big)^3\Big[n(3n^3+71n^2+411n+635)\Big(\frac{S_1}{n}\Big)^3\\
&\qquad + 3(29n^3+273n^2+533n+285)\Big(\frac{S_1}{n}\Big)\Big]\\
&+\frac{54n^2}{(n+5)(n+3)(n+1)(n-1)}\Big(\frac{S_2}{n}\Big)^4\Big[(4n^3+43n^2+118n+115)\Big(\frac{S_1}{n}\Big)\Big]\Big\}\Big]. \qquad (5{\cdot}2)
\end{aligned}
$$

Since the expressions appearing with λ_5, $\lambda_3\lambda_4$ and λ_3^3 in (5·2) are odd functions in S_1, the integral of $g_2(S_1, S_2)$ with respect to S_1 will be independent of these cumulants, so that the distribution of s^2 will be the same as that obtained in (3·2) and (3·3).

For the distribution of t we proceed as before using (4·1), and obtain after simplification ($p(t)$ being given by (4·2))

$$q(t) = p(t) + \lambda_5 \frac{\{[n^2+(n-1)^2]\,t^5 + 10(n-1)\,t^3 - 15(n-1)^2\,t\}}{40[\sqrt{(2n\pi)}]\,n(n-1)^2\,[1+t^2/(n-1)]^{\frac{1}{2}(n+5)}}$$

$$+ \lambda_3\lambda_4 \frac{\left\{\begin{array}{l}(40n^3+258n^2+182n+51)\,t^7 - 3(150n^3+896n^2+869n+297)\,t^5 \\ \quad +15(n-1)\,(68n^2+311n+237)\,t^3 - 15(n-1)^2\,(25n+59)\,t\end{array}\right\}}{144[\sqrt{(2\pi n)}]\,n(n-1)^3\,[1+t^2/(n-1)]^{\frac{1}{2}(n+7)}}$$

$$+ \lambda_3^3 \frac{\left\{\begin{array}{l}(64n^4+952n^3+3578n^2+3011n+996)\,t^9 - 9(144n^4+1918n^3+7011n^2 \\ \quad +7550n+2913)\,t^7 + 27(n-1)\,(252n^3+2821n^2+8554n+7077)\,t^5 \\ \quad -45(n-1)^2\,(541n^2+2086n+3049)\,t^3 + 945(n-1)^3\,(3n+13)\,t\end{array}\right\}}{1296[\sqrt{(2\pi n)}]\,n(n-1)^4\,[1+t^2/(n-1)]^{\frac{1}{2}(n+9)}}$$

$$= p_0(t) + \lambda_3 p_{\lambda_3}(t) - \lambda_4 p_{\lambda_4}(t) + \lambda_3^2 p_{\lambda_3^2}(t) + \lambda_5 p_{\lambda_5}(t) + \lambda_3\lambda_4 p_{\lambda_3\lambda_4}(t) + \lambda_3^3 p_{\lambda_3^3}(t). \qquad (5{\cdot}3)$$

Polynomials in t appearing in the numerator of the corrective functions are analogous in form to the Hermitian polynomials of the same degree. The form of the corrective term due to λ_5 is simple but those due to $\lambda_3\lambda_4$ and λ_3^3 are rather complicated.

6. TAIL-AREA PROBABILITIES

We are interested only in the tail area of these corrective terms and so consider the integrals

$$\int_{-\infty}^{-t_0} p(t)\,dt \quad \text{or} \quad \int_{t_0}^{\infty} p(t)\,dt.$$

For the first of the four right-hand members of (4·2) the integral has been tabulated. For the next member the two tail areas will be equal in magnitude but opposite in sign. In each of the other two members we shall have equal areas at both tails. These integrals may be tabulated for given values of n and t_0, either directly from their algebraic expressions or more conveniently by using the Incomplete Beta function table of Pearson. For the sake of completeness we shall give here the expressions for tail areas of all the components, introducing $P_0(t_0)$, $P_{\lambda_3}(t_0)$, $P_{\lambda_4}(t_0)$ and $P_{\lambda_3^2}(t_0)$ to represent them. Thus, for one tail, the corrected probability integral $P(t_0)$, will be given by

$$P(t_0) = P_0(t_0) + \lambda_3 P_{\lambda_3}(t_0) - \lambda_4 P_{\lambda_4}(t_0) + \lambda_3^2 P_{\lambda_3^2}(t_0), \qquad (6{\cdot}1)$$

in which

$$P_0(t_0) = \int_{-\infty}^{-t_0} p_0(t)\,dt = \int_{t_0}^{\infty} p_0(t)\,dt = \tfrac{1}{2} I_{u_0}\left(\frac{n-1}{2}, \frac{1}{2}\right), \qquad (6{\cdot}2)$$

$$P_{\lambda_3}(t_0) = \int_{-\infty}^{-t_0} p_{\lambda_3}(t)\,dt = -\int_{t_0}^{\infty} p_{\lambda_3}(t)\,dt = \frac{1}{6\sqrt{(2n\pi)}} \frac{\left\{1+\dfrac{(2n-1)}{(n-1)}\,t_0^2\right\}}{[1+t_0^2/(n-1)]^{\frac{1}{2}(n+1)}}, \qquad (6{\cdot}3)$$

which is evidently of order $n^{-\frac{1}{2}}$. For calculation, however, it is more convenient to write

$$P_{\lambda_3}(t_0) = \left\{\frac{(2n-1)}{6\sqrt{(2n\pi)}}\right\} I_{u_0}\left(\frac{n-1}{2}, 1\right) - \left\{\frac{n-1}{3\sqrt{(2n\pi)}}\right\} I_{u_0}\left(\frac{n+1}{2}, 1\right). \qquad (6{\cdot}3\,bis)$$

We also have

$$P_{\lambda_4}(t_0) = \int_{-\infty}^{-t_0} p_{\lambda_4}(t)\,dt = \int_{t_0}^{\infty} p_{\lambda_4}(t)\,dt = \left\{\frac{1}{12(n-1)\sqrt{[\pi(n-1)]}} \frac{\Gamma(\frac{1}{2}n)}{\Gamma[\frac{1}{2}(n-1)]}\right\} \frac{t_0^3 - \dfrac{3(n-1)}{(n+1)}\,t_0}{[1+t_0^2/(n-1)]^{\frac{1}{2}(n+2)}} \qquad (6{\cdot}4)$$

$$= \frac{(n-1)}{24} I_{u_0}\left(\frac{n-1}{2}, \frac{1}{2}\right) - \frac{(n-1)\,(n+2)}{12n} I_{u_0}\left(\frac{n+1}{2}, \frac{1}{2}\right) + \frac{(n+4)\,(n-1)}{24n} I_{u_0}\left(\frac{n+3}{2}, \frac{1}{2}\right), \qquad (6{\cdot}4\,bis)$$

and lastly,

$$P_{\lambda_3^2}(t_0) = \int_{-\infty}^{-t_0} p_{\lambda_3^2}(t)\,dt = \int_{t_0}^{\infty} p_{\lambda_3^2}(t)\,dt = \left\{\frac{(2n+5)}{36(n-1)^2}\frac{\Gamma(\frac{1}{2}n)}{\sqrt{[\pi(n-1)]}\,\Gamma[\frac{1}{2}(n-1)]}\right\} \times \frac{t_0^5 + \frac{2(2n-7)(n-1)}{(n+1)(2n+5)}t_0^3 - \frac{3(2n+11)(n-1)^2}{(n+1)(n+3)(2n+5)}t_0}{[1+t_0^2/(n-1)]^{\frac{1}{2}(n+4)}} \tag{6·5}$$

$$= \left[\left\{\frac{(n-1)(2n+5)}{72}\right\} I_{u_0}\left(\frac{n-1}{2}, \frac{1}{2}\right) - \left\{\frac{(n-1)(2n^2+5n+8)}{24n}\right\} I_{u_0}\left(\frac{n+1}{2}, \frac{1}{2}\right) + \left\{\frac{(n-1)(2n^2+5n+12)}{24n}\right\} I_{u_0}\left(\frac{n+3}{2}, \frac{1}{2}\right) - \left\{\frac{(n-1)(2n^2+5n+12)}{72n}\right\} I_{u_0}\left(\frac{n+5}{2}, \frac{1}{2}\right)\right], \tag{6·5 bis}$$

where t_0 is any typical value of t, and the transformed variate u_0 is given by

$$u_0 = \frac{1}{1+t_0^2/(n-1)}$$

and $I_{u_0}(\frac{1}{2}\nu_1, \frac{1}{2}\nu_2)$, is the Incomplete B-function as defined by Pearson. Obviously $P_{\lambda_4}(t_0)$ and $P_{\lambda_3^2}(t_0)$ are of order n^{-1}. For $t_0 = 0(0{\cdot}5)\,4$ and $n = 2, 3, 4, 5, 6, 7, 9, 13, 25$ and ∞, values of $P_0(t_0)$, $P_{\lambda_3}(t_0)$, $P_{\lambda_4}(t_0)$ and $P_{\lambda_3^2}(t_0)$ have been tabulated (Table 1) with the help of Pearson's Incomplete B-function table. The sample sizes are chosen to correspond to the degrees of freedom 1, 2, 3, 4, 5, 6, 8, 12, 24 and ∞, as in Fisher's table of z.

It will be noticed that the effect of λ_4 is very small, as has been inferred by various authors. The corrective tail area beyond $|t_0| = 3$ (which is nearly the 4 and 3 % points in normal samples of 5 and 6 respectively) is at most 1 % when the size of sample does not exceed 5, about 0·3 % for samples of 10 and about 0·1 % for samples of 20. The actual probability is obtained by multiplying these tail areas by the given value of λ_4, and they are additive to that of the normal theory t.

As Geary points out, the effect of λ_3 is rather serious, but so is that of λ_3^2. For a two-sided test the positive and negative tail areas will, of course, balance each other, but those due to λ_3^2 will not.

For a comparative study of the form of the frequency curves of the normal theory t and those of the corrective functions for λ_3, λ_4 and λ_3^2, diagrams have been constructed for $n = 3$, 6, 13 and 25. They are shown in Figs. 1–4. As is to be expected, the curves of the corrective functions tend to those of the third, fourth and sixth derivatives of the normal function as n tends to infinity.

An example to illustrate the use of the tables has been taken from Neyman & Pearson's (1928) paper 'On the use and interpretation of certain test criteria'. The given values of λ_3 and λ_4 are not large and more or less satisfy A. L. Bowley's (1928) criterion for moderately abnormal curves.

Example (Neyman & Pearson, 1928, p. 203). 'Records of weight in a large population of mice for males between 120 and 140 days of age show a mean value of 23·823 g., and for the frequency distribution $\beta_1 = 0{\cdot}086$, $\beta_2 = 2{\cdot}687$. Can the group of six mice with the weights 22·5, 26·0, 20·5, 24·0, 18·0 and 24·5 be considered a random sample from this population?'

Here $\lambda_3 = +0{\cdot}293$ (inferred as positive), $\lambda_3^2 = 0{\cdot}086$, $\lambda_4 = -0{\cdot}313$, and we find $t = -1{\cdot}038$ $(= -t_0$, say) for $n' = 5$ degrees of freedom. Considering both tails, the normal theory probability is

$$P_0(|t| \geqslant t_0) = 0{\cdot}3468,$$

Table 1. *Comparative values of $P_0(t_0)$, $P_{\lambda_3}(t_0)$, $P_{\lambda_4}(t_0)$ and $P_{\lambda_3^2}(t_0)$ for different degrees of freedom n' $(= n-1$, n being the size of sample$)$. $P(t_0) = P_0(t_0) + \lambda_3 P_{\lambda_3}(t_0) - \lambda_4 P_{\lambda_4}(t_0) + \lambda_3^2 P_{\lambda_3^2}(t_0)$*

t_0	$P_0(t)$	$P_{\lambda_3}(t_0)$	$P_{\lambda_4}(t_0)$	$P_{\lambda_3^2}(t_0)$	$P_0(t_0)$	$P_{\lambda_3}(t_0)$	$P_{\lambda_4}(t_0)$	$P_{\lambda_3^2}(t_0)$
	$n' = 1$				$n' = 2$			
0·0	0·5000	0·0470	0·0000	0·0000	0·5000	0·0384	0·0000	0·0000
0·5	0·3524	0·0589	−0·0064	−0·0066	0·3333	0·0495	−0·0069	−0·0066
1·0	0·2500	0·0665	0·0000	0·0044	0·2113	0·0597	−0·0027	0·0009
1·5	0·1872	0·0622	0·0047	0·0147	0·1362	0·0563	0·0025	0·0118
2·0	0·1476	0·0547	0·0064	0·0188	0·0918	0·0469	0·0047	0·0172
2·5	0·1211	0·0476	0·0066	0·0195	0·0648	0·0375	0·0051	0·0179
3·0	0·1024	0·0416	0·0064	0·0188	0·0477	0·0298	0·0047	0·0165
3·5	0·0886	0·0368	0·0059	0·0176	0·0364	0·0239	0·0041	0·0145
4·0	0·0780	0·0329	0·0055	0·0163	0·0286	0·0194	0·0035	0·0125
	$n' = 3$				$n' = 4$			
0·0	0·5000	0·0332	0·0000	0·0000	0·5000	0·0297	0·0000	0·0000
0·5	0·3257	0·0431	−0·0062	−0·0056	0·3217	0·0387	−0·0055	−0·0047
1·0	0·1955	0·0540	−0·0034	−0·0002	0·1870	0·0495	−0·0036	−0·0005
1·5	0·1153	0·0513	0·0013	0·0098	0·1040	0·0473	0·0006	0·0084
2·0	0·0697	0·0413	0·0035	0·0152	0·0581	0·0372	0·0028	0·0135
2·5	0·0439	0·0310	0·0039	0·0157	0·0334	0·0266	0·0031	0·0139
3·0	0·0288	0·0229	0·0034	0·0139	0·0200	0·0184	0·0027	0·0119
3·5	0·0197	0·0169	0·0028	0·0114	0·0124	0·0127	0·0021	0·0095
4·0	0·0137	0·0126	0·0023	0·0093	0·0081	0·0088	0·0016	0·0072
	$n' = 5$				$n' = 6$			
0·0	0·5000	0·0271	0·0000	0·0000	0·5000	0·0251	0·0000	0·0000
0·5	0·3192	0·0355	−0·0049	−0·0041	0·3174	0·0329	−0·0044	−0·0035
1·0	0·1816	0·0397	−0·0035	−0·0005	0·1780	0·0430	−0·0033	−0·0005
1·5	0·0970	0·0440	0·0002	0·0074	0·0921	0·0413	0·0000	0·0066
2·0	0·0510	0·0340	0·0022	0·0122	0·0462	0·0315	0·0019	0·0111
2·5	0·0272	0·0234	0·0025	0·0125	0·0233	0·0210	0·0021	0·0113
3·0	0·0150	0·0154	0·0021	0·0104	0·0120	0·0132	0·0017	0·0092
3·5	0·0086	0·0099	0·0016	0·0079	0·0064	0·0081	0·0012	0·0067
4·0	0·0052	0·0065	0·0011	0·0057	0·0036	0·0050	0·0008	0·0047
	$n' = 8$				$n' = 12$			
0·0	0·5000	0·0222	0·0000	0·0000	0·5000	0·0184	0·0000	0·0000
0·5	0·3153	0·0291	−0·0037	−0·0028	0·3131	0·0243	−0·0027	−0·0019
1·0	0·1733	0·0384	−0·0030	−0·0005	0·1685	0·0325	−0·0023	−0·0002
1·5	0·0860	0·0371	−0·0002	−0·0055	0·0797	0·0315	−0·0004	0·0042
2·0	0·0403	0·0277	0·0014	0·0094	0·0343	0·0230	0·0009	0·0073
2·5	0·0185	0·0177	0·0016	0·0095	0·0140	0·0137	0·0011	0·0072
3·0	0·0085	0·0103	0·0013	0·0074	0·0055	0·0072	0·0008	0·0053
3·5	0·0040	0·0058	0·0008	0·0051	0·0022	0·0036	0·0005	0·0033
4·0	0·0020	0·0032	0·0005	0·0033	0·0009	0·0017	0·0003	0·0019
	$n' = 24$				$n' = \infty$			
0·0	0·5000	0·0133	0·0000	0·0000	0·5000			
0·5	0·3101	0·0176	−0·0015	−0·0010	0·3085			
1·0	0·1636	0·0238	−0·0014	−0·0001	0·1587			
1·5	0·0733	0·0232	−0·0003	0·0025	0·0668			
2·0	0·0285	0·0164	0·0004	0·0043	0·0228			
2·5	0·0098	0·0090	0·0005	0·0041	0·0062			
3·0	0·0031	0·0041	0·0003	0·0028	0·0013			
3·5	0·0009	0·0016	0·0002	0·0015	0·0002			
4·0	0·0003	0·0006	0·0001	0·0007	0·0000			

whereas the estimate of the actual probability

$$P(|t| \geqslant t_0) = 0{\cdot}3447.$$

If we consider the negative tail only, then $P(t \leqslant -t_0) = 0{\cdot}1832$. So that whichever tail is used the conclusion that there is no reason to doubt the origin of the sample is obvious.

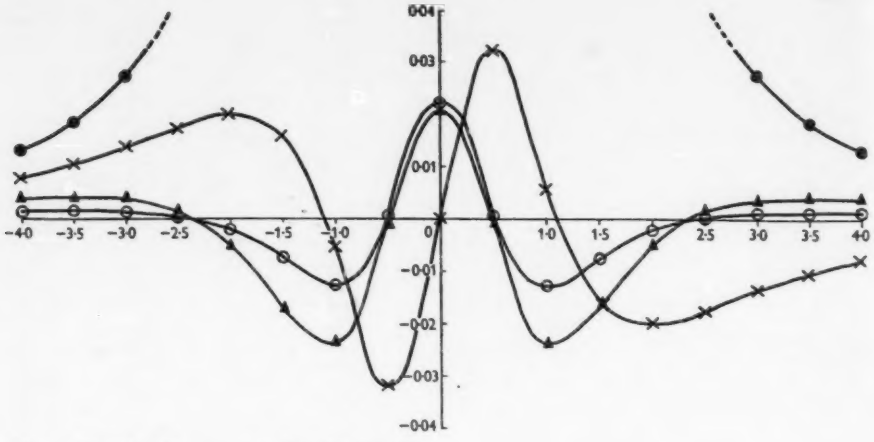

Fig. 1. Showing the terms in the distribution

$p(t) = p_0(t) + \lambda_3 p_{\lambda_3}(t) - \lambda_4 p_{\lambda_4}(t) + \lambda_3^2 p_{\lambda_3^2}(t)$ for $n' = n - 1 = 2$ degrees of freedom.

Explanation of symbols for all figures:

$p_0(t)$; $p_{\lambda_4}(t)$; $p_{\lambda_5}(t)$; $p_{\lambda_5^2}(t)$.

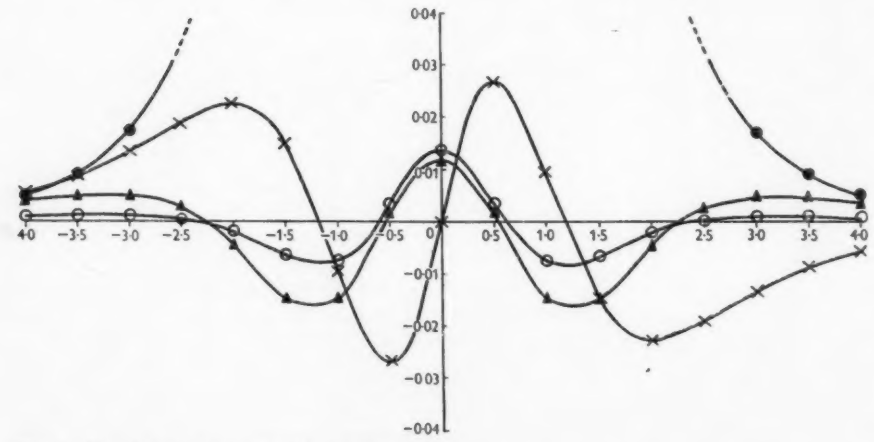

Fig. 2. Showing the terms in the distribution

$p(t) = p_0(t) + \lambda_3 p_{\lambda_3}(t) - \lambda_4 p_{\lambda_4}(t) + \lambda_3^2 p_{\lambda_3^2}(t)$ for $n' = n - 1 = 5$ degrees of freedom.

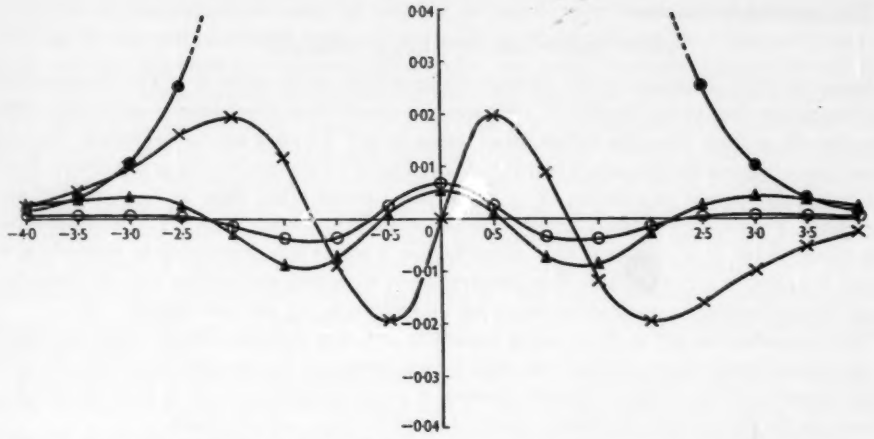

Fig. 3. Showing the terms in the distribution
$p(t) = p_0(t) + \lambda_3 p_{\lambda_3}(t) - \lambda_4 p_{\lambda_4}(t) + \lambda_3^2 p_{\lambda_3^2}(t)$ for $n' = n - 1 = 12$ degrees of freedom.

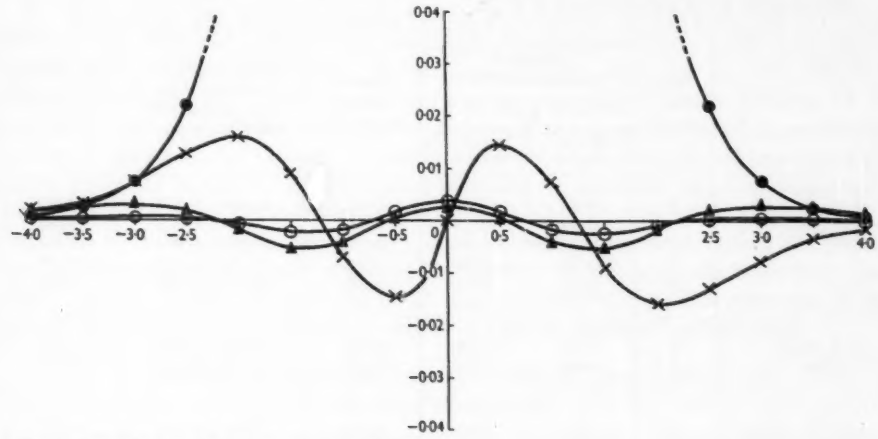

Fig. 4. Showing the terms in the distribution
$p(t) = p_0(t) + \lambda_3 p_{\lambda_3}(t) - \lambda_4 p_{\lambda_4}(t) + \lambda_3^2 p_{\lambda_3^2}(t)$ for $n' = n - 1 = 24$ degrees of freedom.

7. Asymptotic character of the series for $p(t)$

Pointing out some of the asymptotic properties up to order n^{-1} possessed by the expression $p_0(t) + \lambda_3 p_{\lambda_3}(t)$ (being his earlier approximation for 'Student's' t), Geary (1936) observed that for samples of moderate size, the probability

$$\{p_0(t) + \lambda_3 p_{\lambda_3}(t)\}\, dt$$

might have quite an extended range of applicability, provided that at least the lower frequency constants λ_4, λ_5 and λ_6 are small. He considered this to be a matter for experimental investigation.

The probability function

$$p(t) = \{p_0(t) + \lambda_3 p_{\lambda_3}(t) - \lambda_4 p_{\lambda_4}(t) + \lambda_3^2 p_{\lambda_3^2}(t)\}$$

obtained in (4·2) possesses similar asymptotic properties up to order n^{-1}. For it represents the frequency density of 'Student's' t for samples drawn from *any* parent population if the samples are so large that the higher order terms in $n^{-\frac{1}{2}r}$ for $r \geqslant 3$ can be neglected. On the other hand, it gives the frequency density of 'Student's' t for samples of *any size* drawn from the particular parent population (2·1), when the λ terms other than in λ_3, λ_4 and λ_3^2 are negligible. From Geary's (1947) asymptotic formulae for the cumulants of t for any universe (his results 2·18), it is apparent that these higher λ terms can only occur in terms in $n^{-\frac{1}{2}r}$ for $r \geqslant 3$, so that it is not unlikely that for samples of moderate size $p(t)$ has a quite extended range of applicability, provided higher order cumulants λ_5, λ_6, etc., are small.

The expression for $q(t)$ in (5·3), being correct to $n^{-\frac{1}{2}}$, will provide a closer approximation to the actual distribution of t than the other two expressions, for samples from *any* universe if the terms in n^{-2} and higher negative powers of n are negligibly small, as also for samples of *any size* for the parent population specified by (5·1) if $\lambda_6 \ldots \lambda_3^4$ are small.

Table 2. *Showing the experimental determinations of the true probability of t with their standard errors for various sampled universes together with the corresponding probability obtained from the frequency function p(t) of* (4·2)

	Size of sample	Sampled populations with		Normal theory probability	Experimental determinations of the probability with their S.E.'s	Probability as estimated from $p(t)$ of (4·2)
		$\lambda_3^2 = \beta_1$	$\lambda_4 = \beta_2 - 3$			
E. S. Pearson (1929)	5	0·00	−0·50	0·040	0·044 ± 0·009	0·043
	5	0·00	1·12	0·040	0·038 ± 0·009	0·034
	5	0·00	4·07	0·040	0·029 ± 0·008	0·018
	5	0·20	0·30	0·040	0·044 ± 0·009	0·043
	5	0·50	0·73	0·040	0·042 ± 0·009	0·048
H. L. Rietz (1939)	5	0·00	−0·63	0·067	0·058 ± 0·007	0·071
	5	0·03	−0·55	0·067	0·070 ± 0·008	0·071
A. N. K. Nair (1941)	5*	2·00	3·00	0·040	0·062 ± 0·011	0·072
	6†	4·00	6·00	0·030	0·088 ± 0·010	0·088

* Type III curve. † Exponential curve.

Approximate probabilities correct to $n^{-\frac{1}{2}}$ for a single tail will be obtained by applying to the normal theory value necessary corrections due to λ_3, λ_4, λ_3^2 and λ_5, $\lambda_3\lambda_4$, λ_3^3. But when both tails are used, corrections for λ_4 and λ_3^2 only will furnish the same order of approximation since the contributions due to odd-order cumulants cancel each other, i.e. the use of $p(t)$, which provides an approximation correct to n^{-1} only, will actually lead for a two-sided test to results correct to $n^{-\frac{1}{2}}$. This favours the application of our formula for $p(t)$ to samples of moderate size from populations whose degree of departure from normality is not necessarily small.

Results of investigations made by various writers support the conjecture that $p(t)$ or $q(t)$ may have a quite extended range of applicability. Table 2 shows the values of λ_3^2 and λ_4 of the sampled populations, the size of the samples considered, the normal theory probability $P_0(t_0)$, and the experimental determinations with their standard errors, of the approximate probability $P(t_0)$ against their values obtained from $p(t)$. The agreement is quite good, even in cases where samples of five or six only have been taken from markedly non-normal populations. One of the two sampled populations of Nair is a very skew Type III Pearson curve, and the other is the exponential curve. These populations, as is well known, can hardly be represented adequately by the third or fourth approximation to the law of error.

The agreement between experimental and theoretical results, especially for the two populations considered by Nair, may be partly due to the fact that the approximation yielded by $p(t)$ is, as we have noted, not only correct to n^{-1} but also to $n^{-\frac{3}{2}}$, since both tails have been used for these cases.

For certain values of λ_3^2 and λ_4, estimates of $P(t_0)$ at $|t_0| = 3$, for samples of 5 and 6, obtained from formula (4·2), are shown in Tables 3 and 4. Numerical values of λ_3^2 and λ_4 have been selected to include a fairly wide class of non-normal populations met with in practice. In spite of the fact that there is a fair agreement between the tabulated results and the experimental determinations, it is obviously unwise to assume that they necessarily furnish in all cases a satisfactory estimate of the true probability of t for such small samples.

The forms of the two approximations, namely (4·2), correct to n^{-1}, and (5·3), to $n^{-\frac{3}{2}}$, are different from those which may be correspondingly obtained from Geary's (1947) asymptotic expansion to n^{-2} (his result (2·24)). He has obtained that expression by using Charlier's 'Differential Series' with the normal theory t as the generating function, utilizing for the purpose his derived results for the first six cumulants of non-normal t. One of the advantages of Geary's expression for t is that it takes into account some of the higher cumulants and higher powers of λ_3 and λ_4. Accordingly it may have an extended range of applicability. But the fact that it agrees, to order n^{-1}, with M. S. Bartlett's (1935) result suggests that the formula is not satisfactory, since the asymptotic character of the latter cannot be assumed. For Bartlett specified his parent population by the first three terms of the Edgeworth series which take no account of the term in λ_3^2, namely, $10\lambda_3^2\phi^{(6)}(x)/6!$. The estimates of true probabilities shown in Geary's Table 2 cannot as such be regarded as satisfactory.

8. Probability corrections due to higher cumulants of the parent population

For the effects of λ_5, $\lambda_3\lambda_4$, λ_3^3 we consider as before the probability integrals of the corrective functions and denote them by $P_{\lambda_5}(t_0)$, $P_{\lambda_3\lambda_4}(t_0)$ and $P_{\lambda_3^3}(t_0)$ respectively. We then have

$$P_{\lambda_5}(t_0) = -\int_{-\infty}^{-t_0} p_{\lambda_5}(t)\,dt = \int_{t_0}^{\infty} p_{\lambda_5}(t)\,dt$$

$$= \frac{\left\{1 + \frac{2(4n-3)}{(n-1)}t_0^2 + \frac{(2n^2-2n+1)}{(n-1)^2}t_0^4\right\}}{40.n^{\frac{1}{2}}\sqrt{(2\pi)}\left(1+\frac{t_0^2}{n-1}\right)^{\frac{1}{2}(n+3)}} \qquad (8\cdot1)$$

$$= \left\{\frac{(n-1)(n-4)}{20\sqrt{(2\pi)}\,n^{\frac{1}{2}}}\right\} I_{u_0}\left(\frac{n+3}{2}, 1\right) - \left\{\frac{(n-1)(n-2)}{10\sqrt{(2\pi)}\,n^{\frac{1}{2}}}\right\} I_{u_0}\left(\frac{n+1}{2}, 1\right)$$

$$+ \left\{\frac{(2n^2-2n+1)}{40\sqrt{(2\pi)}\,n^{\frac{1}{2}}}\right\} I_{u_0}\left(\frac{n-1}{2}, 1\right), \qquad (8\cdot1\,bis)$$

Table 3. *Showing comparative values of*

$$P_{t_0} = 1 - \int_{-t_0}^{t_0} p(t)\,dt = 2[P_0(t_0) - \lambda_4 P_{\lambda_4}(t_0) + \lambda_3^2 P_{\lambda_3^2}(t_0)],$$

near the 4 % *point (approximately) of normal theory t for samples of* 5 (n' = *degrees of freedom* = 4)

λ_4 \ λ_3^2	0·00	0·20	0·25	0·50	1·0	1·5	2·0	4·6
−2·0	0·0505							
−1·5	0·0479	0·0527	0·0539	0·0598				
−1·0	0·0452	0·0500	0·0512	0·0572	0·0691			
−0·5	0·0426	0·0474	0·0486	0·0545	0·0665	0·0784		
0·0	0·0399	0·0447	0·0459	0·0519	0·0638	0·0758	0·0877	
0·5	0·0373	0·0421	0·0433	0·0492	0·0612	0·0731	0·0851	
1·0	0·0346	0·0394	0·0406	0·0466	0·0585	0·0705	0·0824	
1·5	0·0320	0·0368	0·0380	0·0439	0·0559	0·0678	0·0798	
2·0	0·0293	0·0341	0·0353	0·0413	0·0532	0·0652	0·0771	0·1249
3·0	0·0240	0·0288	0·0300	0·0360	0·0479	0·0599	0·0718	0·1196
4·0	0·0187	0·0235	0·0247	0·0307	0·0426	0·0546	0·0665	0·1143
6·0	0·0081	0·0129	0·0141	0·0201	0·0320	0·0440	0·0559	0·1037

Table 4. *Showing comparative values of*

$$P_{t_0} = 1 - \int_{-t_0}^{t_0} p(t)\,dt = 2[P_0(t_0) - \lambda_4 P_{\lambda_4}(t_0) + \lambda_3^2 P_{\lambda_3^2}(t_0)],$$

near the 3 % *point (approximately) of the normal theory t for samples of* 6 (n' = *degrees of freedom* = 5)

λ_4 \ λ_3^2	0·00	0·20	0·25	0·5	1·0	1·5	2·0	4·0
−2·0	0·0386							
−1·5	0·0365	0·0406	0·0417	0·0468				
−1·0	0·0343	0·0385	0·0393	0·0447	0·0551			
−0·5	0·0322	0·0364	0·0374	0·0426	0·0530	0·0634		
0·0	0·0301	0·0343	0·0353	0·0405	0·0509	0·0613	0·0717	
0·5	0·0280	0·0321	0·0332	0·0384	0·0488	0·0592	0·0696	
1·0	0·0259	0·0300	0·0311	0·0363	0·0467	0·0571	0·0674	
1·5	0·0237	0·0279	0·0289	0·0341	0·0445	0·0549	0·0653	
2·0	0·0216	0·0258	0·0268	0·0320	0·0424	0·0528	0·0632	0·1048
3·0	0·0174	0·0216	0·0226	0·0278	0·0382	0·0486	0·0590	0·1006
4·0	0·0132	0·0173	0·0184	0·0236	0·0340	0·0443	0·0547	0·0963
6·0	0·0047	0·0088	0·0099	0·0151	0·0255	0·0359	0·0463	0·0879

$$P_{\lambda_3\lambda_4}(t_0) = -\int_{-\infty}^{-t_0} p_{\lambda_3\lambda_4}(t)\,dt = \int_{t_0}^{\infty} p_{\lambda_3\lambda_4}(t)\,dt$$

$$= (n_{34})\,I_{u_0}\left(\frac{n-1}{2}, 4\right) - (n_{43})\,I_{u_0}\left(\frac{n+1}{2}, 3\right) + (n_{52})\,I_{u_0}\left(\frac{n+3}{2}, 2\right) - (n_{61})\,I_{u_0}\left(\frac{n+5}{2}, 1\right), \tag{8·2}$$

where (n_{34}), (n_{43}), (n_{52}) and (n_{61}) are given by

$$\left.\begin{aligned}
(n_{34}) &= \left\{\frac{(n-1)\,B\left(\frac{n-1}{2}, 4\right)}{288\sqrt{(2\pi)\,n^{\frac{1}{2}}}}\right\}(40n^3+258n^2+182n+51),\\
(n_{43}) &= \left\{\frac{B\left(\frac{n+1}{2}, 3\right)}{96\sqrt{(2\pi)\,n^{\frac{1}{2}}}}\right\}(150n^3+896n^2+869n+297),\\
(n_{52}) &= \left\{\frac{5B\left(\frac{n+3}{2}, 2\right)}{96\sqrt{(2\pi)\,n^{\frac{1}{2}}}}\right\}(68n^2+311n+237),\\
(n_{61}) &= \left\{\frac{5B\left(\frac{n+5}{2}, 1\right)}{96\sqrt{(2\pi)\,n^{\frac{1}{2}}}}\right\}(25n+59),
\end{aligned}\right\} \tag{8·3}$$

$$P_{\lambda_3^2}(t_0) = \int_{-\infty}^{-t_0} p_{\lambda_3^2}(t)\,dt = -\int_{t_0}^{\infty} p_{\lambda_3^2}(t)\,dt$$

$$= -(n_{45})\,I_{u_0}\left(\frac{n-1}{2}, 5\right) + (n_{54})\,I_{u_0}\left(\frac{n+1}{2}, 4\right) - (n_{63})\,I_{u_0}\left(\frac{n+3}{2}, 3\right)$$

$$+ (n_{72})\,I_{u_0}\left(\frac{n+5}{2}, 2\right) - (n_{81})\,I_{u_0}\left(\frac{n+7}{2}, 1\right), \tag{8·4}$$

where

$$\left.\begin{aligned}
(n_{45}) &= \left\{\frac{(n-1)\,B\left(\frac{n-1}{2}, 5\right)}{2592\sqrt{(2\pi)\,n^{\frac{1}{2}}}}\right\}(64n^4+952n^3+3578n^2+3011n+996),\\
(n_{54}) &= \left\{\frac{B\left(\frac{n+1}{2}, 4\right)}{228\sqrt{(2\pi)\,n^{\frac{1}{2}}}}\right\}(144n^4+1918n^3+7011n^2+7550n+2913),\\
(n_{63}) &= \left\{\frac{B\left(\frac{n+3}{2}, 3\right)}{96\sqrt{(2\pi)\,n^{\frac{1}{2}}}}\right\}(252n^3+2821n^2+8554n+7077),\\
(n_{72}) &= \left\{\frac{5B\left(\frac{n+5}{2}, 2\right)}{288\sqrt{(2\pi)\,n^{\frac{1}{2}}}}\right\}(541n^2+2086n+3049),\\
(n_{81}) &= \left(\frac{35B\left(\frac{n+7}{2}, 1\right)}{96\sqrt{(2\pi)\,n^{\frac{1}{2}}}}\right)(3n+13),
\end{aligned}\right\} \tag{8·5}$$

with the usual notations of Complete and Incomplete B-functions.

The additional formulae (8·1), (8·2) and (8·4) of this section enable us to deduce the probability of t when the parent population can be represented by the fourth approximation. The corrected probability $P(t_0)$, for a single tail, may be obtained from the formula

$$P(t_0) = P_0(t_0) + \lambda_3 P_{\lambda_3}(t_0) - \lambda_4 P_{\lambda_4}(t_0) + \lambda_3^2 P_{\lambda_3^2}(t_0) + \lambda_5 P_{\lambda_5}(t_0) + \lambda_3\lambda_4 P_{\lambda_3\lambda_4}(t_0) + \lambda_3^3 P_{\lambda_3^3}(t_0). \quad (8{\cdot}6)$$

It should be noted that for the negative tail of the distribution $P_{\lambda_3}(t_0)$ and $P_{\lambda_3^2}(t_0)$ are positive, whereas $P_{\lambda_5}(t_0)$ and $P_{\lambda_3\lambda_4}(t_0)$ are negative, the reverse being the case with the other tail.

For samples of 10 from such a population the probability at $t_0 = -2{\cdot}262$ (here $|t_0|$ is the 5 % point of the normal theory t), will be given by

$$P(-2{\cdot}262) = 0{\cdot}0250 + \lambda_3(0{\cdot}0209) - \lambda_4(0{\cdot}0015) + \lambda_3^2(0{\cdot}0092) + \lambda_5(-0{\cdot}0017) + \lambda_3\lambda_4(-0{\cdot}0111) + \lambda_3^3(0{\cdot}0221).$$

If $\lambda_3 = 1$, $\lambda_4 = 1$, then for a Pearson curve, $\lambda_5 = 0{\cdot}25$, and in that case the probability is found to be 0·0642, whereas the estimate from terms as far as λ_3^2 only is 0·0536. Thus for a single tail they differ considerably from the normal theory probability 0·0250.

Summary

The theoretical distribution of 'Student's' t in non-normal samples of any size has been derived with reference to the parent population specified by a number of terms of the Edgeworth series. It contains, in addition to the normal theory frequency function of t, corrective terms due to the cumulants λ_3, λ_4, λ_3^2 and λ_5, $\lambda_3\lambda_4$, λ_3^3. It is assumed that the values of the population λ's are known.

If the population can be well represented by the third or the fourth approximation (in which cases next higher λ's and higher powers of λ_3 and λ_4 are small), then the corresponding t-distributions will be accurate for any size of sample, as far as similar order terms in λ's.

The two expressions for t, namely, the one including the terms up to that in λ_3^2 and the other up to that in λ_3^3, are asymptotic to order n^{-1} and $n^{-\frac{1}{2}}$ respectively, and as such they afford closer approximations to the actual distribution of t for large samples from *any* universe. The probabilities calculated from the former will be the same as those from the latter, if both the tails of t are used. So it is not unlikely that for moderate size of sample, the former may have an extended range of applicability. The satisfactory agreement between the theoretical values and the experimental results of various writers appears to support this conjecture.

Tail-area probabilities of the corrective terms for λ_3, λ_4 and λ_3^2 have been tabulated. An example is given to illustrate the use of the tables. Diagrams showing the frequency curve of non-normal t for certain values of n have been constructed.

The probability corrections due to higher cumulants of the parent population have also been considered for a representative value of n.

I wish to acknowledge my indebtedness to Dr H. E. Daniels for his kind advice and criticism in the course of my investigations, also to Dr J. Wishart for suggesting a number of improvements to the paper. My thanks are due to Mr D. A. East for drawing the diagrams.

REFERENCES

BARTLETT, M. S. (1935). *Proc. Camb. Phil. Soc.* **31**, 223.
BOWLEY, A. L. (1928). *F. Y. Edgeworth's Contributions to Mathematical Statistics.* London: Royal Statistical Society.
CRAMÉR, H. (1928). *Skand. AktuarTidskr.* **11**, 13, 141.
CRAMÉR, H. (1946). *Mathematical Methods of Statistics.* Princeton University Press.
EDGEWORTH, F. Y. (1905). *Trans. Camb. Phil. Soc.* **20**, 36, 113.
EDGEWORTH, F. Y. (1906). *J.R. Statist. Soc.* **69**, 497.
GEARY, R. C. (1936). *J.R. Statist. Soc. Suppl.* **3**, 178.
GEARY, R. C. (1947). *Biometrika*, **34**, 209.
NAIR, A. N. K. (1941). *Sankhyā*, **5**, 393.
NEYMAN, J. & PEARSON, E. S. (1928). *Biometrika*, **20A**, 175, 263.
PEARSON, E. S. (1928, 1929). *Biometrika*, **20A**, 356 and **21**, 259.
RIETZ, H. L. (1939). *Ann. Math. Statist.* **10**, 265.

THE COMBINATION OF PROBABILITIES ARISING FROM DATA IN DISCRETE DISTRIBUTIONS

By H. O. LANCASTER, *Rockefeller Fellow in Medicine*

Introductory

1. Two common methods of combining probabilities from different experiments make use of the additive properties of χ^2. Thus the results of the various experiments may each be expressed as a standardized normal deviate or as its square. In these cases the probabilities are readily combined by the summation of χ^2. In other cases there may be obtained the probability, on the null hypothesis, that a result as divergent as the observed or one more divergent would occur. Such probabilities may be combined in a simple way by the use of the transformation $\chi^2 = -2\log_e P$. No difficulties are met in data from continuous populations nor in discrete populations, where many different observations are possible, such as occur with large samples. However, with discrete populations, as, for example, the binomial with low index or the fourfold table with small numbers, biases arise which may diminish the power of the tests. This is most easily seen in the case of the binomial if independent values of χ^2_c, that is, χ^2 corrected for continuity, are summed. As these biases have not previously received much attention, we examine the problem in some detail. We suggest that, for the χ^2 with one degree of freedom, only the crude χ^2 is suitable for summation. No obvious modification for the probability integral transformation is available, but we suggest the use of the 'mean value χ^2' or an approximation to it, the 'median value χ^2'. The former, when the null hypothesis is true; has an expectation rigorously equal to the theoretical value of 2 and a variance slightly less than the theoretical value of 4.

We discuss, first, the case of the binomial with low index, where an enumeration of all possible events and their relative frequencies of occurrence under the null or other hypothesis is practicable. Then we pass on to the more difficult case of the fourfold tables where the number of possible events is greatly increased, even with the simplified conditions which we have selected.

The binomial

2. *Notation.* In the discussion of the binomial, we shall take p and q to be the probability of success or failure of an individual observation as specified by the null hypothesis, n the number of trials, m the observed number of successes, and p' and q' the corresponding probabilities in the actual population which is being sampled. We shall use P in the usual sense of the calculated probability of the observation and all observations more extreme, and P' to be the probability of these more extreme observations alone, so that $(P-P')$ is the probability (i.e. the relative frequency of occurrence) of the observed result. Occasionally, as in (1), it will be necessary to use P as a continuous variable. In connexion with the probability integral transformation, we define the mean value χ^2, (χ^2_m), and the median value χ^2, (χ'^2_m), as follows:

$$\chi^2_m = \int_{P'}^{P} (-2\log_e P)\,dP/(P-P')$$
$$= 2 - 2\{P\log_e P - P'\log_e P'\}/(P-P'), \tag{1}$$

$$\left.\begin{aligned}\chi'^2_m &= -2\log_e \tfrac{1}{2}\{P+P'\}, && \text{if } P' \neq 0,\\ &= 2-2\log_e P, && \text{if } P'=0.\end{aligned}\right\} \tag{2}$$

3. *A note on the correction for continuity in the case of the binomial distribution.* Yates (1934) made a definite, although indirect, statement on the use of his continuity correction for the combination of probabilities from different experiments in the following words:

'$P(\chi')$ [i.e. our $P(\chi_c)$] gives 0·1427, an excellent approximation, whereas $P(\chi)$ gives 0·0612, which, though not in itself attaining significance, is less than half its true value; this would be exceedingly misleading if a number of such probabilities from different classes of experiment were to be combined.' D. Mainland (1948) takes this to infer that χ_c^2, and not the crude χ^2, is to be summed, a fallacious practice already rejected by Cochran (1942). We may note that

$$\begin{aligned} E(\chi_c^2) &= E(|\, m-np\,|-\tfrac{1}{2})^2/(npq) \\ &= E(m-np)^2/(npq) - E\,|\, m-np\,|/(npq) + 1/(4npq) \\ &= 1 + 1/(4npq) - (\text{mean deviation})/(\text{variance}). \end{aligned} \tag{3}$$

If np is integral, the expectation is less than that given above by a quantity equal to the frequency of the modal term divided by four times the variance. We have always here diminished $|\, m-np\,|$ by $\frac{1}{2}$. This will result in a change of sign if $|\, m-np\,|$ is less than $\frac{1}{2}$, but, in general, it is obvious that only a trivial and non-systematic difference will be made to the expectation by the alternative method of not diminishing $|\, m-np\,|$ by $\frac{1}{2}$ if it be less than $\frac{1}{2}$, and the method used here has rendered the algebra much easier. If neither n, np nor nqp be small, this bias due to the use of the continuity correction is trivial, but then the continuity correction would be unnecessary; in those cases for which the correction is said to be essential, the mean deviation may be of the same order of magnitude as the variance or, at any rate, not a small fraction of it, and so bias will result. We note, for example, that in the binomial $(\frac{1}{2}+\frac{1}{2})^2$ the expectation of χ_c^2 is 0·25 if the null hypothesis be true. Many will regard this as an unnecessarily extreme case, but it is not as extreme as Mainland's example (on his p. 54) which we quote in § 9 below.

4. *The power of the methods of combination of probabilities.* In Table 1 A we have attempted to give some idea of the power of the tests, in the sense of Neyman & Pearson (1933), for a binomial population with index of 5. We note first of all that no single experiment can give a result significant at the 5 % level, so that we are led to some method of combination of successive experiments, and we have compared the results of (*a*) simple pooling of the successes and failures, (*b*) summation of the crude χ^2, (*c*) summation of the mean value χ^2, (*d*) summation of the $-2\log_e P$, (*e*) summation of χ_c^2, which we find are decreasingly powerful in that order.* We have taken $p = 0{\cdot}5$ as the null hypothesis and have enumerated the possible events in populations with various values of p' and have calculated the relative frequency of each event. We have then been able to calculate the frequency with which we may expect results significant at the 5 % level combining two, three or four experiments; in the rows indicated by the symbol ∞, we have shown the limits to which results tend if very many samples are combined. Thus, we have been able to show that, if the null hypothesis were true in the population sampled, by combining two experiments it would be rejected in 4·3 % of cases using the summation of the crude χ^2, whereas if the p' of the population sampled were 0·25 or 0·75, then it would be rejected in some 25·2 % of cases and so on. If a χ^2 has an expectation below its appropriate theoretical value of either 1 or 2, respectively, then we

* In all cases we have used a two-sided test, e.g. where the null hypothesis gives the binomial $(\frac{1}{2}+\frac{1}{2})^5$, we have pooled together results with 0 or 5, 1 or 4 and 2 or 3 successes.

may assume that in an indefinitely large number of experiments the value of χ^2 obtained is below that of the 5 % significance level in practically every set of experiments combined, and so we have assumed that no significant results will be obtained in these cases. However, if the expectation is above the theoretical value, then we have assumed that a summation of the χ^2 from an indefinitely large number of experiments combined will always yield a significant result. Thus we find that, if $p' = 0{\cdot}33$, the expected value of $(-2\log_e P)$ is 1·480, and so, if the experiment were to be repeated a thousand times we should have a χ^2 of approximately 1480 with 2000 degrees of freedom and should have no ground for rejecting the null hypothesis. In this example we have scored the rejection rate as zero. If, on the other hand, the expectation of χ^2 were above 2, as in the case where $p' = 0{\cdot}2$, we should expect every indefinitely long series of experiments to lead to the rejection of the null hypothesis. As was to be anticipated, simple pooling was found to be the most powerful method of combination. The crude χ^2 and χ^2_m were equally powerful in this particular group of populations, whereas χ^2_c and the probability integral transformation were much less powerful and, in fact, were unable to detect considerable departures from the null hypothesis even with a large number of experiments available. Somewhat similar findings occur with samples from the populations $(p'+q')^{10}$ when the null hypothesis specifies that $p = 0{\cdot}5$, as can be seen from Table 1B. In these discrete populations we can only give arithmetical results in specific cases, since the use of approximate methods would remove the effect of the discreteness which is what we desire to study. We have spent some time on the binomial because the problems here are more easily studied than in the case of the fourfold table where the enumeration of cases becomes impracticable, and because the binomial is the limiting case of the series involved in the fourfold table, when the members of one row become indefinitely large.

5. *The probability integral transformation*, $\chi^2 = -2\log_e P$. We discuss now in greater detail the transformation, $\chi^2 = -2\log_e P$, as introduced by R. A. Fisher (1932). A similar transformation was dealt with by Karl Pearson (1933) and later E. S. Pearson (1938). None of these authors appears to have considered the effects of discontinuity, but the recent increasing use of very small samples renders such an inquiry now necessary. This transformation gives unbiased results when based on continuous distributions, but it will be shown that it must be used with caution in the case of discrete distributions. Under the null hypothesis, all values of P are considered to be equally likely, so that, if F is the distribution function of P,

$$dF = dP. \tag{4}$$

Integrating, this gives

$$F = P \quad (0 < P \leqslant 1), \tag{5}$$

since the probability of obtaining $P = 0$ is 0. Then $f(P)$, the frequency function, is given by

$$f(P) = \text{constant} = 1 \quad (0 < P \leqslant 1). \tag{6}$$

After the transformation, $y = -2\log_e P$, we find that y has a frequency function

$$f(y) = \tfrac{1}{2}\exp(-\tfrac{1}{2}y), \tag{7}$$

so that y is distributed as χ^2 with 2 degrees of freedom. Thus, any number of probabilities, P, may be converted by means of this transformation to χ^2 for 2 degrees of freedom, and, using the additive properties of the χ^2 distribution, the χ^2 may be summed together with the degrees of freedom in the usual manner. It is necessary, however, to note that P must be

Table 1 A

Table 1A

The table shows the results of drawing samples of five from a population where the probability of a single success is p' and of combining the results of these drawings by different techniques. In each case the hypothesis tested is that p equals 0·5 and the significance level is 5 %. The body of the table shows (i) the power of the test, that is, the percentage frequency with which the null hypothesis will be rejected using a 5 % significance level, and (ii) the expectations of the various χ^2.

Method of testing	No. of experiments	Method of combination	Value of p' in the population tested						Criterion of significance at 5 % level
			0·5	0·4 or 0·6	0·33 or 0·66	0·3 or 0·7	0·25 or 0·75	0·2 or 0·8	0·1 or 0·9	
			(i) *Percentage of significant results*							
Exact method	1	—	0	0	0	0	0	0	0	No significant results possible
	2	(*a*) Simple pooling	2·15	4·80	10·44	14·94	24·41	37·58	73·61	0, 1, 9, 10 successes
	3	,,	3·52	9·24	20·95	29·70	46·13	64·82	94·44	0 3, 12–15 successes
	4	,,	4·14	12·72	29·74	41·64	61·72	80·86	98·86	0–5, 15–20 successes
	∞	,,	5·00	100·00	100·00	100·00	100·00	100·00	100·00	p not exactly 0·5
χ_c^2	2	(*c*) Summation of χ_c^2	0·39	0·77	1·84	2·91	5·68	10·75	34·87	$\Sigma\chi_c^2 > 5{\cdot}991$
	3	,,	0·02	0·07	0·25	0·50	1·35	13·53	20·59	$\Sigma\chi_c^2 > 7{\cdot}815$
	4	,,	0·09	0·25	0·90	1·73	4·44	10·64	45·89	$\Sigma\chi_c^2 > 9{\cdot}488$
	∞	,,	0	0	0	0	100·00	100·00	100·00	$E(\chi_c^2) > 1$*
$\chi^2 = -2\log_e P$	2	(*d*) Summation of $-2\log_e P$	0·39	0·77	1·84	2·91	5·68	10·75	34·87	$\Sigma\chi^2 > 9{\cdot}488$
	3	,,	0·39	0·85	2·30	3·88	8·34	16·96	54·95	$\Sigma\chi^2 > 12{\cdot}592$
	4	,,	0·09	0·25	0·90	1·73	4·44	10·64	45·89	$\Sigma\chi^2 > 15{\cdot}507$
	∞	,,	0	0	0	0	100·00	100·00	100·00	$E(\chi^2) > 2$*
Crude χ^2†	2	(*b*) Summation of χ^2	4·30	6·69	11·90	16·15	25·20	38·05	73·66	$\Sigma\chi^2 > 5{\cdot}991$
	3	,,	2·95	5·17	10·62	15·45	26·35	42·25	82·54	$\Sigma\chi^2 > 7{\cdot}815$
	4	,,	2·92	5·45	11·92	17·73	30·79	49·24	95·15	$\Sigma\chi^2 > 9{\cdot}488$
	∞	,,	5·00	100·00	100·00	100·00	100·00	100·00	100·00	$E(\chi^2) > 1$*
			(ii) *Expectations of the various* χ^2							
χ_c^2			0·450	0·550	0·731	0·856	1·091	1·382	2·152	
Crude χ^2			1·000	1·160	1·144	1·640	2·000	2·440	3·560	
$\chi^2 = -2\log_e P$			0·960	1·147	1·480	1·708	2·126	2·635	3·919	
χ_m^2			2·000	2·233	2·633	2·910	3·418	4·035	5·588	
$\chi_m'^2$			1·890	2·117	2·521	2·789	3·308	3·931	5·515	

* E indicates the expectation.

† χ_m^2 and $\chi_m'^2$ give the same results as the crude χ^2 for the combination of 2 or 3 experiments.

Table 1B

The table illustrates the problem considered in Table 1A, except that it now relates to the results of drawing samples of $n = 10$ and combining them in various ways. The null hypothesis is that $p = 0{\cdot}5$.

Method of testing	No. of experiments	Method of combination	Value of p' in the population tested					Criterion of significance at 5 % level
			0·5	0·45 or 0·55	0·4 or 0·6	0·33 or 0·66	0·25 or 0·75	
				(i) *Percentage of significant results*				
Exact method	1	—	2·15	2·78	4·80	10·44	24·41	0, 1, 9, 10 successes
	2	(*a*) Simple pooling	4·14	6·18	12·72	29·74	61·72	0–5, 15–20 successes
	3	,,	4·28	7·44	17·72	43·18	80·35	0–9, 21–30 successes
	∞	,,	5·00	100·00	100·00	100·00	100·00	p not exactly 0·5
χ_c^2	2	(*c*) Summation of χ_c^2	0·77	1·13	2·50	7·65	25·06	$\Sigma\chi_c^2 > 5{\cdot}991$
	3	,,	0·98	1·47	3·41	14·05	37·14	$\Sigma\chi_c^2 > 7{\cdot}815$
	∞	,,	0	0	0	100·00	100·00	$E(\chi_c^2) > 1$*
$\chi^2 = -2\log_e P$	2	(*d*) Summation of $-2\log_e P$	1·69	2·24	4·66	12·46	34·57	$\Sigma\chi^2 > 9{\cdot}488$
	3	,,	0·98	1·47	3·41	14·05	37·14	$\Sigma\chi^2 > 12{\cdot}592$
	∞	,,	0	0	0	100·00	100·00	$E(\chi^2) > 2$*
Crude χ^2	2	(*b*) Summation of χ^2	5·02	6·46	11·11	23·72	50·80	$\Sigma\chi^2 > 5{\cdot}991$
	3	,,	4·40	6·09	11·31	27·57	62·34	$\Sigma\chi^2 > 7{\cdot}815$
	∞	,,	5·00	100·00	100·00	100·00	100·00	$E(\chi^2) > 1$*
				(ii) *Expectations of the various* χ^2				
χ_c^2			0·583	0·650	0·852	1·340	2·325	
Crude χ^2			1·000	1·090	1·360	2·000	3·250	
$\chi^2 = -2\log_e P$			1·296	1·406	1·735	2·513	4·025	
χ_m^2			2·000	2·128	2·511	3·407	5·125	
$\chi_m'^2$			1·934	2·059	2·431	3·304	4·985	

Notes. All methods give the same percentage of significant results in a single experiment.
χ_m^2 gives the same results as the crude χ^2 for combination of two and three experiments.
$\chi_m'^2$ gives a slightly lower percentage of significant results than the crude χ^2 and χ_m^2.

* E indicates the expectation.

assumed to come from a rectangular population with end-points 0 and 1. Further, in continuous populations the transforms of P and $(1-P)$ both have a finite expectation. This is not so with discrete populations as can easily be seen in the case of the binomial, since there is a finite probability of obtaining a result $1-P=0$, when $-2\log_e(1-P)=\infty$. If, moreover, a two-sided comparison is made so that we always sum the probabilities from the tail in which the observation lies, a further difficulty arises. In order to define which tail to use, we must fix some dividing point and sum from the lower tail if the observation falls below this point and otherwise sum from the upper tail. In the case of a continuous distributon the most natural point to take is the median of the sampling distribution and then we may take for P in the transformation, double the probability integral calculated from the appropriate tail. This was the procedure suggested by Sukhatme (1935). Mainland (1948) has suggested using such a method when dealing with discrete distributions, but without some adjustment it may lead to values of P greater than unity. It is worth studying some examples to appreciate the difficulties more fully.

6. *Examples of some difficulties arising in the transformation,* $\chi^2=-2\log_e P$. (*a*) The expectation of χ^2 is always below 2 for discrete populations when the null hypothesis is true, as we now illustrate in the case of the binomial $(\frac{1}{2}+\frac{1}{2})^4$. We may tabulate the possible events, their relative frequency and the corresponding values of P and $(-2\log_e P)$ for a *one-sided comparison*:

Successes (m)	Relative frequency of m	Cumulative probability (P)	$-2\log_e P$
4	0·0625	1·0000	0
3	0·2500	0·9375	0·12908
2	0·3750	0·6875	0·74939
1	0·2500	0·3125	2·32630
0	0·0625	0·0625	5·54518

We conclude that the expectation of χ^2, i.e. of $-2\log_e P$, is 1·241, when the null hypothesis, that $p=0{\cdot}5$, is true.

For a *two-sided comparison*, we obtain the results set out below:

Successes (m)	Probability of (m)	Cumulative probability (P)	$-2\log_e P$
0 or 4	0·125	0·125	4·15888
1 or 3	0·500	0·625	0·94007
2	0·375	1·000	0·00000

Here, the expectation of χ^2 is 0·990 instead of the theoretical value of 2. It will be noted that in this case the division point is at $m=2$, and the procedure adopted is equivalent to summing from the lower tail on half the occasions when $m=2$ and from the upper tail in the other half and, then, taking P as double this tail sum.

Notes. All methods give the same percentage of significant results in a single experiment.
χ^2_m gives the same results as the crude χ^2 for combination of two and three experiments.
χ^2_m gives a slightly lower percentage of significant results than the crude χ^2 and χ^2_m.
* *E* indicates the expectation.

(*b*) When, however, the distribution is asymmetric, a further modification is needed. Consider the case of the binomial $(\frac{2}{3}+\frac{1}{3})^5$; here the mean lies between $m = 1$ and $m = 2$. Col. 3 of the table shows the cumulated probabilities from either tail. In the final column, instead of multiplying by 2, we have adjusted P by dividing by the proportion of observations in the respective tail. This avoids the possibility of having a probability, P, more than unity.

Successes (m)	Probability of m	Cumulative probability, P from one tail	Adjusted P
0	0·131687	0·131687	0·28570
1	0·329218	0·460905	1·00000
...	...	...	...
2	0·329218	0·539094	1·00000
3	0·164609	0·209876	0·38931
4	0·041152	0·045267	0·08397
5	0·004115	0·004115	0·00763

(*c*) As another illustration, take the binomial $(\frac{2}{3}+\frac{1}{3})^6$. Here the question arises—how should the observations corresponding to a value $m = 2$ (the mean) be apportioned? Following the method of the preceding example, we should get different adjusted values for P and therefore different expectations for $-2\log_e P$, according to the tail to which we assign samples with $m = 2$. It would, of course, be possible, but troublesome in practice, to obtain a median division by using some random method of allocating samples having $m = 2$, to the lower tail with a probability of 0·14884 and to the upper tail with a probability of 0·18038.

Successes (m)	Probability of m	Cumulative probability (P) summed from appropriate tail
0	0·08779	0·08779
1	0·26337	0·35116
2	0·32922	?
3	0·21948	0·31962
4	0·08231	0·10014
5	0·01646	0·01783
6	0·00137	0·00137

7. *The mean value* χ^2. To illustrate the solution proposed, suppose we are dealing with samples of 5, that the null hypothesis is that $p = \frac{1}{3}$ and that we are using a one-sided test, i.e. only concerned with the alternative that $p' > \frac{1}{3}$. The following table shows the relevant quantities. On the null hypothesis, $y = -2\log_e P$ can assume the six discrete values given in col. 4 with the corresponding relative frequencies given in col. 2. In Fig. 1 the latter have been plotted as ordinates against the former as abscissae. If P were a continuous variable, then $y = -2\log_e P$ would have the exponential frequency function $f(y)$ of equation (7); this curve also is shown in the diagram. The mean of the frequency curve is at 2. The mean of the

discrete distribution (i.e. the sum of the products of the corresponding numbers in cols. 2 and 4 of the table) is 1·314. There is clearly a considerable bias.

Case of binomial $(\frac{2}{3}+\frac{1}{3})^5$; *one-sided test*

i	Relative frequency $P_i - P_{i+1}$	Cumulative probability P_i	$y_i = -2\log_e P_i$	$\bar{y}_i$	$-2\log_e \frac{1}{2}(P_i + P_{i+1})$
(1)	(2)	(3)	(4)	(5)	(6)
0	0·131687	1·000000	0·0000	0·1379	0·1362
1	0·329218	0·868313	0·2824	0·7213	0·7028
2	0·329218	0·539095	1·2357	2·0329	1·9644
3	0·164609	0·209877	3·1225	4·2788	4·1181
4	0·041153	0·045268	6·1903	7·7108	7·4026
5	0·004115	0·004115	10·9862	12·9862	12·9862

	Expectation	Variance
χ^2 with 2 D.F. (theoretical)	2·000	4·000
$\bar{y}_i$	1·314	2·482
$\overline{y_i}$	2·000	3·689
$-2\log_e \frac{1}{2}(P_i + P_{i+1})$	1·932	3·443

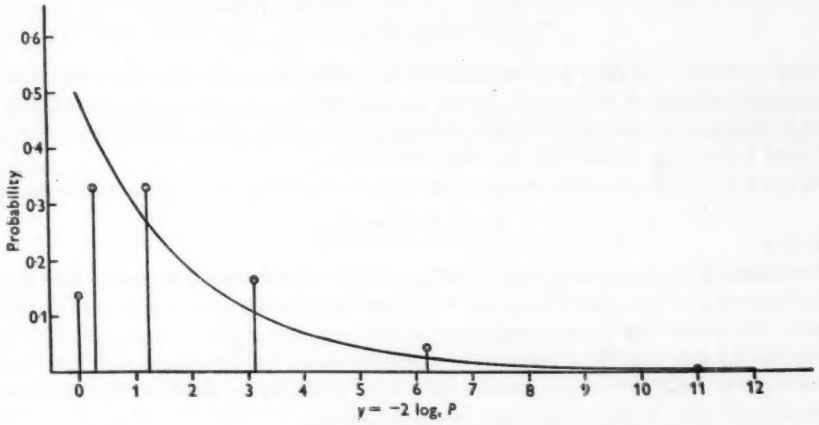

Fig. 1. (*a*) Frequency curve $f(y) = \frac{1}{2}e^{-\frac{1}{2}y}$ and (*b*) probabilities for the binomial $(\frac{2}{3}+\frac{1}{3})^5$.

In general, suppose that there are possible s successive values of $-2\log_e P$ and that we denote these by

$$y_i = -2\log_e P_i \quad (i = 0, 1, 2, \ldots, s-1). \tag{8}$$

Call $P_0 = 1$, $y_0 = 0$, $P_s = 0$ and $y_s = \infty$. Then

$$P_i - P_{i+1} = \int_{y_i}^{y_{i+1}} f(y)\,dy. \tag{9}$$

Thus the bias in the expectation of y_i arises because the frequency which in the continuous

distribution is spread between y_i and y_{i+1}, is concentrated in the discrete distribution at the lower end of the interval. To correct this, we propose to replace y_i in the test by

$$\bar{y}_i = \int_{y_i}^{y_{i+1}} yf(y)\,dy \Big/ \int_{y_i}^{y_{i+1}} f(y)\,dy \tag{10}$$

$$= \int_{P_{i+1}}^{P_i} (-2\log_e P)\,dP/(P_i - P_{i+1}) = 2 - 2\{P_i \log_e P_i - P_{i+1}\log_e P_{i+1}\}/(P_i - P_{i+1}). \tag{11}$$

It will be seen from equation (10) that we are replacing $y_i = -2\log_e P_i$ by the mean value of y in the interval y_i to y_{i+1}, as given by the frequency function $f(y)$ appropriate for the continuous variable case. By doing this we ensure that the expectation of $\bar{y}_i$ is 2, for

$$E(\bar{y}_i) = \sum_{i=0}^{s-1} (P_i - P_{i+1})\bar{y}_i = \sum_{i=0}^{s-1} \left\{ \int_{P_{i+1}}^{P_i} (-2\log_e P)\,dP \right\} = \int_0^1 (-2\log_e P)\,dP = 2. \tag{12}$$

Further, the variance of $\bar{y}_i$ will approach 4 from below, as the number of groups is increased. To show this, we note that the variance of the frequency function, i.e. of the distribution of χ^2 with 2 degrees of freedom, may be split up as follows:

$$\begin{aligned}\int_0^\infty f(y)\,(y-2)^2\,dy &= \sum_{i=0}^{s-1} \left\{ \int_{y_i}^{y_{i+1}} f(y)\,(y-2)^2\,dy \right\} \\ &= \sum_{i=0}^{s-1} \left\{ \int_{y_i}^{y_{i+1}} f(y)\,(\bar{y}_i - 2)^2\,dy + \int_{y_i}^{y_{i+1}} f(y)\,(y-\bar{y}_i)^2\,dy \right\} \\ &= \sum_{i=0}^{s-1} (P_i - P_{i+1})\,(\bar{y}_i - 2)^2 + \sum_{i=0}^{s-1} \left\{ \int_{y_i}^{y_{i+1}} f(y)\,(y-\bar{y}_i)^2\,dy \right\}. \end{aligned} \tag{13}$$

The first expression on the right-hand side of (13) is the variance of $\bar{y}_i$, which therefore falls short of the variance of χ^2, i.e. of 4, by the second term on the right-hand side of (13). This latter expression depends on the width of intervals of y imposed by the null hypothesis and will usually be small, unless there are very few groups.

We see from (11) that, in the event of the observation being the most extreme possible,

$$\bar{y}_{s-1} = 2 - 2\log_e P_{s-1}, \tag{14}$$

since $P_s = 0$.

If we write χ^2_m for $\bar{y}_i$, P for P_i and P' for P_{i+1}, we have the form of the mean value χ^2 given at the beginning of this paper in equation (1). The distribution of χ^2_m, or $\bar{y}_i$, is of course still discrete, but since it has an expectation of 2 and a variance slightly under 4, it may be expected that if k independent samples are combined by summing the values of χ^2_m, the resulting statistic will be distributed approximately as χ^2 with $2k$ degrees of freedom. The bias involved in summing $-2\log_e P$ has, at any rate, been eliminated.

In order to avoid the tedious calculations necessary for χ^2_m, we have defined

$$\chi'^2_m = -2\log_e \tfrac{1}{2}(P_i + P_{i+1}) = -2\log_e \tfrac{1}{2}(P + P') \tag{15}$$

for all values of P', except when $P' = 0$. In this case, χ^2_m has the simple form given by equation (14), and we assign this value also to χ'^2_m. This slightly reduces the bias of χ'^2_m, which has an expectation a little below 2. It is found, for the ordinary sets of probabilities which may arise in practice, that χ^2_m and χ'^2_m approximate closely. Their expectations for a number of cases are given in Tables 1 A and 1 B, and are seen to be approximately equal. For the example considered at the beginning of the present section, the individual values of $\chi^2_m = \bar{y}_i$ and $\chi'^2_m = -2\log_e \tfrac{1}{2}(P_i + P_{i+1})$ are shown in cols. (5) and (6) of the table on p. 377; expectations and variances are compared below the table.

THE FOURFOLD TABLE

8. In considering the problem for a fourfold table, we shall illustrate the position on the case where it is supposed that samples of ten are drawn randomly from two populations in which the chances of an individual possessing a certain character are $p = (1-q)$ and $p' = (1-q')$, respectively. The sample results will therefore be recorded as shown below. On the assumption that $p = p'$, and within any set of samples for which $a+b$ = constant, a follows a distribution of hypergeometric type with

$$E(a) = \tfrac{1}{2}(a+b), \quad \text{Variance of } a = (a+b)(20-a-b)/76.$$

	No. with character	No. without	Totals
1st sample	a	$10-a$	10
2nd sample	b	$10-b$	10
Totals	$a+b$	$20-a-b$	20

The definitions of the quantities P, P', χ_m^2, $\chi_m'^2$ given in § 2 for the binomial can therefore be extended to apply to the hypergeometric. Further, for the crude χ^2 and χ_c^2, we have

$$\left.\begin{aligned} \chi^2 &= 19(a-b)^2/\{(a+b)(20-a-b)\}, \\ \chi_c^2 &= 20(|a-b|-1)^2/\{(a+b)(20-a-b)\}. \end{aligned}\right\} \tag{16}$$

To compute $-2\log_e P$ we calculate the sum of the hypergeomtric terms in the usual way, assuming $a+b$ fixed, and make a two-tailed comparison. We note that under the null hypothesis, $p = p'$,

$$E(\chi^2) = 1, \tag{17}$$

$$E(\chi_c^2) = 20\{1-(\text{mean deviation}-\tfrac{1}{2})/(\text{variance})\}/19. \tag{18}$$

If $p \neq p'$, the expected frequencies of a and b will be given by the terms of the product of the two binomials, $(p+q)^{10}$ and $(p'+q')^{10}$. Thus

$$P(a,b \mid p,p') = {}^{10}C_a p^a q^{10-a} \times {}^{10}C_b p'^b q'^{10-b}. \tag{19}$$

Using (19) we can compute, for given p and p', the probability or relative frequency of each fourfold table that can arise and hence obtain the expectations of χ^2, χ_c^2, etc. Results so obtained are given in Table 2. It will be seen that the expectation of χ_c^2, and also that of the probability integral transformation $(-2\log_e P)$, is always less than the theoretical value, when the null hypothesis is true. The advantage of using χ rather than χ_c in repeated sampling was discussed from a rather different angle in connexion with the fourfold table by E. S. Pearson (1947, pp. 151–6). Both the expectations of χ_c^2 and $-2\log_e P$ are also less than the theoretical value when there is some moderate departure from the null hypothesis. The crude χ^2 as defined by (16) and the χ_m^2 both have an expectation of the theoretical magnitude when the null hypothesis is true and an expectation above the theoretical, 1 or 2 respectively, when it is false. In the case of the binomial, we were able to compare the effect of using various methods of combination by an enumeration of cases and a calculation of the probabilities of obtaining a result significant at the 5 % level, when 2, 3, 4 and a very large number of

Table 2. *A table to show the expectations of the various χ^2 calculated from the fourfold tables as formed by the methods of this article*

Null hypothesis	Value of p	Value of p'	Expectations of			
			χ_c^2	Crude χ^2	$-2\log_e P$	χ_m^2
True	0·5	0·5	0·482	1·000	1·040	2·000
	0·45	0·45	0·480	1·000	1·035	2·000
	0·40	0·40	0·471	1·000	1·021	2·000
	0·33	0·33	0·448	1·000	0·984	2·000
	0·25	0·25	0·387	1·000	0·904	2·000
False	0·55	0·25	1·715	2·738	2·989	4·292
	0·50	0·25	1·326	2·229	2·408	3·631
	0·45	0·25	1·009	1·805	1·906	3·080
	0·60	0·40	1·009	1·751	1·904	3·007
	0·55	0·45	0·612	1·188	1·257	2·255

experiments are combined. In the case of the fourfold table the computational work would be prohibitive, so we have calculated a number k such that $k \times E(\chi^2)$ is greater than the 5 % point of the tabulated χ^2 for the appropriate number of degrees of freedom. In other words

Table 3

Value of p	Value of p'	Values of k for different methods of combination			
		Crude χ^2	$-2\log_e P$	Mean value χ^2	χ_c^2
0·55	0·25	3	14	3	14
0·50	0·25	5	100	6	60
0·45	0·25	11	*	12	10,000
0·60	0·40	13	*	13	10,000
0·55	0·45	200 (approx.)	*	300 (approx.)	*

* In these cases the test cannot be said to have power.

we are calculating how many experiments we shall have to combine to give a power to our test of 50 %, assuming that in a number of experiments added together in this way there is a probability of 0·5 of obtaining a total greater than its expectation. The results are tabulated in Table 3. Thus, for example, when $p = 0{\cdot}55$, $p' = 0{\cdot}25$, we obtain from Table 2:

k	$k \times E(-2\log_e P) = k \times 2{\cdot}989$	Degrees of freedom ($2k$)	5 % point of χ^2 for $2k$ degrees of freedom
13	38·857	26	38·885
14	41·846	28	41·337

Thus, we conclude that there is not a 50 % chance of establishing significance at the 5 % level, using $-2\log_e P$, until fourteen samples are available for combination. On the other hand, for these same values of p and p', using crude χ^2 and combining three samples we find that $3 \times E(\chi^2) = 3 \times 2{\cdot}738 = 8{\cdot}214$, and is greater than the 5 % significance level of χ^2 for 3 degrees of freedom, namely, 7·815.

It is to be noted that the crude χ^2 is slightly more powerful than the mean value χ^2_m, but that both are very much superior to χ^2_c or to the χ^2 obtained from the usual form of the probability integral transformation. Where the expectation of χ^2 in the fourfold tables is below the theoretical value, we have assigned no value to k since it is clear that an indefinite repetition of the experiment would not add to the power of the test.

9. *An instructive example.* The following fourfold table is taken from Mainland (1948, p. 54). The animals are from his 'group A'.

	Died	Survived	Total
Treated animals	0	12	12
Control animals	2	10	12
Total	2	22	24

We find that in the set of samples with these margins, the exact probability that $a = 0$ on the null hypothesis is 11/46 or 0·2391304, and the corresponding value for χ^2_c is $6/11 = 0{\cdot}5455$. For all possible tables with this set of marginal totals, $E(\chi^2_c)$ is only 0·2609 if the null hypothesis be true. Mainland, however, comments thus: 'Again it appears that the treatment tends to reduce mortality, but in neither group, with $P = 0{\cdot}025$ as the standard, is the difference significant. Group A, $n = 12$, $P = 0{\cdot}2391$.' On the next page he goes on to use this value of χ^2_c and a similar one from group B to obtain a χ^2 for two degrees of freedom. We may note that the observation recorded is the furthest possible from the null hypothesis, and yet, were we to obtain it repeatedly, we should soon obtain a significantly low value of χ^2. Such usages cannot be defended. In the example above, $-2\log_e P$ has an expectation of only 0·9694 for the two-tailed test in place of the theoretical 2. However, we could calculate a mean value χ^2_m which will have an expectation of exactly 2 on the null hypothesis.

10. *An advantage of the probability integral transformation.* We have seen that the probability integral transformation, using χ^2_m or χ'^2_m of equations (1) and (2) has approximately equal power with the crude χ^2, when used in the summation of two way tests from the binomial distribution and fourfold tables; it has, however, the additional advantage that it can also be used for the combination of one-way tests and thus derive additional power in those cases where the departures from the null hypothesis we are looking out for are in the one direction. We can see that this is equivalent to the addition of a quantity $(-2\log_e \frac{1}{2})$, or 1·39 approximately, to the χ^2_m or χ'^2_m of every experiment which gives a probability near zero. Of course, similar transformations could be found whereby the probability from an experiment can be transformed to the scale of χ^2 for any number of degrees of freedom, but the distributions of χ^2 for 1 and 2 degrees have obvious advantages. First, the data are often already available in the scale of χ or of χ^2 for 1 degree of freedom; secondly, the simplicity

of the probability integral transformation is a very real advantage. Further, the suggested mean value χ^2_m allows significance tests to be carried out and interpreted when the summed χ^2_m is low. Such tests cannot be carried out with the ordinary $\chi^2 = -2\log_e P$ transformation, since it has an expectation below the theoretical value, even when the null hypothesis is true.

Summary

11. Some numerical investigations have been carried out in the case of the binomial distribution and for the fourfold table. It is shown for the combination of significance tests that the usual probability integral transformation test loses much information and a slight modification is suggested. It appears evident that there is no justification for the use of χ^2_c in the combination of experiments. For combination of two-tailed significance tests, the crude χ^2 is the next most powerful test after simple pooling, and it can be readily used in the combination of fourfold tables where pooling may not be possible. When the use of crude χ^2 is not practicable, it is suggested that the mean-value form of the probability integral transformation χ^2_m, defined in equation (1), should be used. A simple approximation to χ^2_m is derived and given in equation (2); this we have termed the median value χ^2, χ'^2_m. For the combination of probability tests when it is desirable to pay regard to the direction of the observed deviation, either the mean or median value probability integral transformation is superior to the crude χ^2.

This work was carried out while the author was a Rockefeller Fellow in Medicine. He would like to thank Prof. A. Bradford Hill of the London School of Hygiene for the facilities of his department, Dr J. O. Irwin and Mr P. Armitage for help in clearing up doubtful points, and to Prof. E. S. Pearson for advice and help in redrafting the article.

REFERENCES

Cochran, W. G. (1942). The χ^2 correction for continuity. *Iowa St. Coll. J. Sci.* **16**, 421.

Fisher, R. A. (1932). *Statistical Methods for Research Workers*, 5th ed. Edinburgh: Oliver and Boyd.

Mainland, D. (1948). Statistical methods in medical research. I. Qualitative statistics (enumeration data). *Canad. J. Res.* **26**, 1.

Neyman, J. & Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypothesis. *Philos. Trans.* A, **231**, 289.

Pearson, E. S. (1938). The probability integral transformation for testing goodness of fit and combining independent tests of significance. *Biometrika*, **30**, 134.

Pearson, E. S. (1947). The choice of statistical tests illustrated on the interpretation of data classed in a 2×2 table. *Biometrika*, **34**, 139.

Pearson, K. (1933). On a method of determining whether a sample of size n supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random. *Biometrika*, **25**, 379.

Sukhatme, P. V. (1935). *Proc. Ind. Acad. Sci.* **2**, 584.

Yates, F. (1934). Contingency tables involving small numbers and the χ^2 test. *J. Roy. Statist. Soc. Suppl.* **1**, 217.

NOTE ON THE APPLICATION OF FISHER'S *k*-STATISTICS

By F. N. DAVID

1. The *k*-statistics were introduced by R. A. Fisher in 1928. In a paper, remarkable not only for the brilliance of the statistical technique but also for the condensation of the mathematical argument, he set out the essential properties of the *k*-statistics, demonstrated how powers and product moments and cumulants could be evaluated, and gave basic tables for their use. Subsequent development has been carried out by J. Wishart (1929*a*, *b*, 1930, 1933) and M. G. Kendall (1940*a*, *b*, *c*, 1942). Wishart worked chiefly on the multivariate case, while M. G. Kendall concerned himself principally with proving rigorously the rules of procedure set out by Fisher. In the course of his analysis he gave certain methods which enable some of the algebra involved in Fisher's original procedure to be cut short. Nevertheless, even with these short cuts the algebra necessary is heavy, and it appeared worth while to investigate the possibility of shortening it still further, particularly for those cases where it is not possible to make the κ_1 of the parent population equal to zero.

2. The technique involved in the application of the *k*-statistics to any problem of distribution may be summarized briefly in the following way. We suppose that it is desired to find the sampling moments of some statistic involving powers and/or products of the one-part symmetric functions

$$s_r = \sum_{i=1}^{n} x_i^r \quad (r = 1, 2, 3, \ldots).$$

The statistic is written in the form of the *k*-statistics by direct substitution. Using Fisher's notation we write

$$\mathscr{E}(k_1^\alpha k_2^\beta k_3^\gamma k_4^\delta) = \mu(1^\alpha 2^\beta 3^\gamma 4^\delta).$$

This product moment, which should properly be written as $\mu'(1^\alpha 2^\beta 3^\gamma 4^\delta)$, is connected with the product cumulants through the identity

$$\kappa(r)\frac{t_1}{1!} + \kappa(s)\frac{t_2}{1!} + \kappa(r^2)\frac{t_1^2}{2!} + \kappa(rs)\frac{t_1}{1!}\frac{t_2}{1!} + \kappa(s^2)\frac{t_2^2}{2!} + \ldots$$

$$\equiv \log\left(1 + \mu'(r)\frac{t_1}{1!} + \mu'(s)\frac{t_2}{1!} + \mu'(r^2)\frac{t_1^2}{2!} + \mu'(rs)\frac{t_1}{1!}\frac{t_2}{1!} + \mu'(s^2)\frac{t_2^2}{2!} + \ldots\right),$$

the extension to any number of variables being obvious.

Fisher's basic tables give expressions for the product cumulants in terms of the parent population cumulants and hence by substitution $\mu'(1^\alpha 2^\beta 3^\gamma 4^\delta)$ may be evaluated.

3. The relations between the product moments and the product cumulants can most easily be derived by making use of an ingenious process due to Kendall (1940*c*), which considerably simplifies the labour. Since this method is not widely known we illustrate it here for fourth powers and two *k*-statistics. We start from the well-known relation that for the parent population

$$\mu_4' = \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4,$$

and hence for the fourth power of the *k*-statistics of order r we may write

$$\mu'(r^4) = \kappa(r^4) + 4\kappa(r^3)\kappa(r) + 3\kappa^2(r^2) + 6\kappa(r^2)\kappa^2(r) + \kappa^4(r).$$

Operating on $\mu'(r^4)$ by $s\partial/\partial r$, and cancelling the factor 4, we have

$$\mu'(r^3s) = \kappa(r^3s) + 3\kappa(r^2s)\,\kappa(r) + \kappa(r^3)\,\kappa(s) + 3\kappa(rs)\,\kappa(r^2) + 3\kappa(r^2)\,\kappa(r)\,\kappa(s) + 3\kappa(rs)\,\kappa^2(r) + \kappa^3(r)\,\kappa(s),$$

where $$\kappa(r) = \kappa_r \quad \text{and} \quad \kappa(s) = \kappa_s,$$

the parent population cumulants of order r and s. Again operating we have

$$\mu'(r^2s^2) = \kappa(r^2s^2) + 2\kappa(rs^2)\,\kappa(r) + 2\kappa(r^2s)\,\kappa(s) + \kappa(s^2)\,\kappa(r^2) + \kappa(s^2)\,\kappa^2(r) + \kappa(r^2)\,\kappa^2(s) + 2\kappa^2(rs) + 4\kappa(rs)\,\kappa(r)\,\kappa(s) + \kappa^2(r)\,\kappa^2(s),$$

results which can be checked from the fundamental identity given above. For three variables we operate on the appropriate power of r by $s\partial/\partial r$ and $t\partial/\partial r$ and so on.

4. It is clear that in order to evaluate $\mu'(r^hs^l)$ when $h+l=4$ or more, a great deal of elementary algebra is required. It is suggested, therefore, that attention might be focused on

$$\mu(r^hs^l) = \mathscr{E}(k_r-\kappa_r)^h\,(k_s-\kappa_s)^l.$$

It is easy to show, using the relations of § 3, that for $h+l=3$ or less

$$\mu(r^hs^l) = \kappa(r^hs^l),$$

as might be expected. For $h+l=4$ we have

$$\begin{aligned}
\mu(r^4) &= \mathscr{E}(k_r-\kappa_r)^4 &&= \kappa(r^4) + 3\kappa^2(r^2),\\
\mu(r^3s) &= \mathscr{E}(k_r-\kappa_r)^3\,(k_s-\kappa_s) &&= \kappa(r^3s) + 3\kappa(rs)\,\kappa(r^2),\\
\mu(r^2s^2) &= \mathscr{E}(k_r-\kappa_r)^2\,(k_s-\kappa_s)^2 &&= \kappa(r^2s^2) + \kappa(s^2)\,\kappa(r^2) + 2\kappa^2(rs).
\end{aligned}$$

The algebra involved by direct substitution is, however, a little heavy, and it is simpler to note that

$$\mathscr{E}(k_r-\kappa_r)^3\,(k_s-\kappa_s) \quad \text{and} \quad \mathscr{E}(k_r-\kappa_r)^2\,(k_s-\kappa_s)^2$$

can be derived from $\mu_4 = \kappa_4 + 3\kappa_2^2$ and $\mu(r^4) = \kappa(r^4) + 3\kappa^2(r^2)$

by Kendall's operative process just described. The relations for $h+l=5$, 6 may thus easily be shown to be

$$\begin{aligned}
\mathscr{E}(k_r-\kappa_r)^5 = \mu(r^5) &= \kappa(r^5) + 10\kappa(r^3)\,\kappa(r^2),\\
\mu(r^4s) &= \kappa(r^4s) + 6\kappa(r^2s)\,\kappa(r^2) + 4\kappa(r^3)\,\kappa(rs),\\
\mu(r^3s^2) &= \kappa(r^3s^2) + 6\kappa(r^2s)\,\kappa(rs) + 3\kappa(rs^2)\,\kappa(r^2) + \kappa(r^3)\,\kappa(s^2),\\
\mathscr{E}(k_r-\kappa_r)^6 = \mu(r^6) &= \kappa(r^6) + 15\kappa(r^4)\,\kappa(r^2) + 10\kappa^2(r^3) + 15\kappa^3(r^2),\\
\mu(r^5s) &= \kappa(r^5s) + 5\kappa(r^4)\,\kappa(rs) + 10\kappa(r^3s)\,\kappa(r^2) + 10\kappa(r^3)\,\kappa(r^2s) + \kappa^2(r^2)\,\kappa(rs),\\
\mu(r^4s^2) &= \kappa(r^4s^2) + \kappa(r^4)\,\kappa(s^2) + 8\kappa(r^3s)\,\kappa(rs) + 6\kappa(r^2s^2)\,\kappa(r^2) + 6\kappa^2(r^2s)\\
&\qquad + 4\kappa(r^3)\,\kappa(rs^2) + 12\kappa(r^2)\,\kappa^2(rs) + 3\kappa^2(r^2)\,\kappa(s^2),\\
\mu(r^3s^3) &= \kappa(r^3s^3) + 3\kappa(r^3s)\,\kappa(s^2) + 9\kappa(r^2s^2)\,\kappa(rs) + 3\kappa(rs^3)\,\kappa(r^2) + 9\kappa(r^2s)\,\kappa(rs^2)\\
&\qquad + \kappa(r^3)\,\kappa(s^3) + 6\kappa^3(rs) + 9\kappa(r^2)\,\kappa(s^2)\,\kappa(rs).
\end{aligned}$$

5. In order to avoid confusion with Fisher's notation we shall write

$$\mu(r^hs^l) = \mathscr{E}(k_r-\kappa_r)^h\,(k_s-\kappa_s)^l = (\overline{r^hs^l}).$$

It would seem that if a table of $(\overline{r^h s^l})$ functions were available a certain amount of the algebra involved in the application of k-statistic technique would be avoided. For example, suppose we consider the almost classical example of finding the moments of

$$x = \sqrt{\left(\frac{(n-1)(n-2)}{6n}\right)} k_3 k_2^{-\frac{3}{2}}$$

in samples of n drawn from a normal population and take the process of obtaining $\mathscr{E}(x^2)$. We have

$$\mathscr{E}(x^2) = \frac{(n-1)(n-2)}{6n}\mathscr{E}((k_3^2)(k_2)^{-3}) = \frac{(n-1)(n-2)}{6n\kappa_2^3}\mathscr{E}\left((k_3^2)\left(1+\frac{k_2-\kappa_2}{\kappa_2}\right)^{-3}\right).$$

The method as outlined by Fisher (1928) and quoted by Kendall (1943) consists in expanding the right-hand bracket and taking expectations

$$\mathscr{E}(x^2) = \frac{(n-1)(n-2)}{6n\kappa_2^3}\left[\mu'(3^2) - \frac{3}{\kappa_2}\mu'(3^2 2) + \frac{6}{\kappa_2^2}\mu'(3^2 2^2) - \frac{10}{\kappa_2^3}\mu'(3^2 2^3) + \ldots\right].$$

The product moments are next put in terms of the product cumulants $\kappa(r^h s^l)$, whence, remembering that k_2 is measured from its expected value, we have

$$\mathscr{E}(x^2) = \frac{(n-1)(n-2)}{6n\kappa_2^3}\left[\kappa(3^2) - \frac{3}{\kappa_2}\kappa(3^2 2) + \frac{6}{\kappa_2^2}(\kappa(3^2 2^2) + \kappa(3^2)\kappa(2^2)) - \frac{10}{\kappa_2^3}(\kappa(3^2 2^3) + 3\kappa(3^2 2)\kappa(2^2) + \kappa(3^2)\kappa(2^3)) + \ldots\right].$$

Reference to Fisher's tables gives the product cumulants in terms of the parent population cumulants and the process is complete. As an alternative procedure suppose we write

$$\mathscr{E}(x^2) = \frac{(n-1)(n-2)\kappa_3^2}{6n\kappa_2^3}\mathscr{E}\left(\left(1+\frac{k_3-\kappa_3}{\kappa_3}\right)^2\left(1+\frac{k_2-\kappa_2}{\kappa_2}\right)^{-3}\right).$$

Upon expansion and taking expectations we have immediately the $(\overline{r^h s^l})$ functions, and all that remains is to substitute from tables of these functions to obtain $\mathscr{E}(x^2)$ at once in terms of the parent population cumulants. The saving in heavy algebra is considerable, even though the original expansion is made a little more clumsy.*

6. A complete table of $(\overline{r^h s^l})$ functions is obviously not necessary. For $h+l \leqslant 3$ the expressions for these functions are the product cumulants already tabled by Fisher or by Kendall. When $h+l>3$ a corrective term must be added to the product cumulant. Table 1 gives a selection of these corrective terms, expressed as functions of the parent-population cumulants, which have been found useful in certain recent investigations. It will be noted for use in deciding order in expansions that product cumulants of weight r will be of order $1/n^{r-1}$. The corrective terms to a product cumulant of weight r will be of order $1/n^{r-2}$ if the corrective terms are two product cumulants multiplied together, of order $1/n^{r-3}$, where they are three and so on. Thus for $h+l=6$ we have $\kappa(r^h s^l)$ is of order $1/n^5$, $\kappa(r^m)\kappa(r^n s^l)$ is of order $1/n^4$ and $\kappa^3(r^p)$ is of order $1/n^3$. If, therefore, no terms are required beyond those involving $1/n^3$ it

* For symmetrical populations it is of course simpler to note that k_2 and k_3 are uncorrelated, and to turn the product of the expectations into $(r^h s^l)$ functions.

will only be necessary to enumerate a single corrective term as the contribution from $\kappa(r^h s^l)$ $(h+l=6)$. In Table 1, printed on p. 392 below, the complete corrective terms are given for $h+l=4, 5$, and those corrective terms for $h+l=6$ which involve $1/n^3$.

7. The question of the validity of the expansions in § 5 is a troublesome one. I have given elsewhere reasons showing that such a process can be justified in an approximate way for some statistics. For other statistics it is difficult to justify except in a limited range of cases. An example of this can be found in the coefficient of variation. We define this, for a sample of n elements, as

$$v^2 = \frac{s_c^2}{\bar{x}^2},$$

where $\bar{x}$ has its usual meaning and

$$s_c^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2.$$

In samples from a normal population it is clear that no matter what the mean and standard deviation of the parent population, the probability exists that $\bar{x}$ can be close to zero and can be negative, and hence the true distribution of v has, theoretically, infinite moments. In order, therefore, to obtain approximations to the moments of v it will be necessary to truncate the parent population at the point $x=0$. We must then make the further assumption that n is chosen large enough for the proportionate frequency lying outside a given multiple of the standard deviation of the mean and of the standard deviation of the variance to be negligible. Under these two assumptions the expansion will be valid. If the parent population is not normal but is, for example, a Pearson curve with start fixed at zero or at some positive value, then it is necessary to make only the assumption regarding n, the sample size.

8. It is assumed that we have a parent population which has cumulants $\kappa_1, \kappa_2, \ldots$ up to any order required. A sample of n elements is randomly and independently drawn from such a population, and to each element of the sample we attach a random variable $x_1, x_2, \ldots, x_n$. We may then write

$$v = \frac{s_c}{\bar{x}} = \frac{k_2^{\frac{1}{2}}}{k_1} = V\left(1+\frac{k_2-\kappa_2}{\kappa_2}\right)^{\frac{1}{2}}\left(1+\frac{k_1-\kappa_1}{\kappa_1}\right)^{-1},$$

where

$$V = \frac{\kappa_2^{\frac{1}{2}}}{\kappa_1},$$

and is the coefficient of variation in the parent population. Under the restrictions just mentioned we may expand the right-hand side of the expression for v, and, for example, for the second moment about an arbitrary origin we obtain

$$\frac{1}{V^2}\mathscr{E}(v^2) = \mathscr{E}\left(1+\frac{k_2-\kappa_2}{\kappa_2}\right)\left(1-\frac{2(k_1-\kappa_1)}{\kappa_1}+\frac{3(k_1-\kappa_1)^2}{\kappa_1^2}-\frac{4(k_1-\kappa_1)^3}{\kappa_1^3}+\frac{5(k_1-\kappa_1)^4}{\kappa_1^4}-\ldots\right),$$

or in terms of the $(\overline{r^h s^l})$ functions

$$\frac{1}{V^2}\mathscr{E}(v^2) = 1-\frac{2}{\kappa_1}(\overline{1})+\frac{3}{\kappa_1^2}(\overline{1}^2)-\frac{4}{\kappa_1^3}(\overline{1}^3)+\frac{5}{\kappa_1^4}(\overline{1}^4)-\ldots$$
$$+\frac{1}{\kappa_2}(\overline{2})-\frac{2}{\kappa_1\kappa_2}(\overline{2\,1})+\frac{3}{\kappa_1^2\kappa_2}(\overline{2\,1^2})-\frac{4}{\kappa_1^3\kappa_2}(\overline{2\,1^3})+\ldots.$$

It is obvious that $(\overline{1})$ and $(\overline{2})$ are zero, and the other terms may be written down from tables of product cumulants and Table 1 of corrective terms.

9. The first four sampling moments of the distribution of v are, to order $1/n^2$,

$$\frac{1}{V}\mathscr{E}(v) = \mu_1'\left(\frac{v}{V}\right) = 1+\frac{1}{n}\left(-\frac{1}{8}\frac{\kappa_4}{\kappa_2^2}-\frac{1}{2}\frac{\kappa_3}{\kappa_1\kappa_2}+\frac{\kappa_2}{\kappa_1^2}\right)-\frac{1}{4(n-1)}$$
$$+\frac{1}{n^2}\left(-\frac{15}{128}\frac{\kappa_4^2}{\kappa_2^4}-\frac{3}{16}\frac{\kappa_4\kappa_3}{\kappa_2^3\kappa_1}+\frac{\kappa_6}{16\kappa_2^3}-\frac{1}{4}\frac{\kappa_3^2}{\kappa_2^2\kappa_1^2}+\frac{\kappa_5}{8\kappa_1\kappa_2^2}+\frac{3}{8}\frac{\kappa_4}{\kappa_1^2\kappa_2}-\frac{5}{2}\frac{\kappa_3}{\kappa_1^3}+\frac{3\kappa_2^2}{\kappa_1^4}\right)$$
$$+\frac{1}{n(n-1)}\left(\frac{9}{32}\frac{\kappa_4}{\kappa_2^2}+\frac{1}{8}\frac{\kappa_3}{\kappa_2\kappa_1}-\frac{1}{4}\frac{\kappa_2}{\kappa_1^2}\right)+\frac{1}{(n-1)^2}\left(\frac{1}{4}\frac{\kappa_3^2}{\kappa_2^3}+\frac{1}{32}\right),$$

$$\frac{1}{V^2}\mathscr{E}(v-\mathscr{E}(v))^2 = \mu_2\left(\frac{v}{V}\right) = \frac{1}{n}\left(\frac{1}{4}\frac{\kappa_4}{\kappa_2^2}-\frac{\kappa_3}{\kappa_1\kappa_2}+\frac{\kappa_2}{\kappa_1^2}\right)+\frac{1}{2(n-1)}$$
$$+\frac{1}{n^2}\left(\frac{7}{32}\frac{\kappa_4^2}{\kappa_2^4}+\frac{1}{4}\frac{\kappa_4\kappa_3}{\kappa_2^3\kappa_1}-\frac{\kappa_6}{8\kappa_2^3}+\frac{1}{4}\frac{\kappa_3^2}{\kappa_2^2\kappa_1^2}-\frac{1}{4}\frac{\kappa_5}{\kappa_1\kappa_2^2}+\frac{5}{2}\frac{\kappa_4}{\kappa_1^2\kappa_2}-\frac{10\kappa_3}{\kappa_1^3}+\frac{8\kappa_2^2}{\kappa_1^4}\right)$$
$$-\frac{1}{n(n-1)}\left(\frac{5}{8}\frac{\kappa_4}{\kappa_2^2}+\frac{1}{2}\frac{\kappa_3}{\kappa_2\kappa_1}-\frac{\kappa_2}{\kappa_1^2}\right)-\frac{1}{(n-1)^2}\left(\frac{1}{2}\frac{\kappa_3^2}{\kappa_2^3}+\frac{1}{8}\right),$$

$$\frac{1}{V^3}\mathscr{E}(v-\mathscr{E}(v))^3 = \mu_3\left(\frac{v}{V}\right) = \frac{1}{n^2}\left(-\frac{3}{16}\frac{\kappa_4^2}{\kappa_2^4}+\frac{1}{8}\frac{\kappa_6}{\kappa_2^3}+\frac{9}{4}\frac{\kappa_3^2}{\kappa_2^2\kappa_1^2}-\frac{3}{4}\frac{\kappa_5}{\kappa_2^2\kappa_1}+\frac{3\kappa_4}{\kappa_2\kappa_1^2}-\frac{10\kappa_3}{\kappa_1^3}+\frac{6\kappa_2^2}{\kappa_1^4}\right)$$
$$+\frac{1}{n(n-1)}\left(\frac{3}{4}\frac{\kappa_4}{\kappa_2^2}-\frac{3\kappa_3}{\kappa_2\kappa_1}+\frac{3\kappa_2}{\kappa_1^2}\right)+\frac{1}{(n-1)^2}\left(\frac{1}{2}\frac{\kappa_3^2}{\kappa_2^3}+\frac{1}{4}\right),$$

$$\frac{1}{V^4}\mathscr{E}(v-\mathscr{E}(v))^4 = \mu_4\left(\frac{v}{V}\right) = \frac{1}{n^2}\left(\frac{3}{16}\frac{\kappa_4^2}{\kappa_2^4}-\frac{3}{2}\frac{\kappa_4\kappa_3}{\kappa_2^3\kappa_1}+\frac{3\kappa_3^2}{\kappa_2^2\kappa_1^2}+\frac{3}{2}\frac{\kappa_4}{\kappa_1^2\kappa_2}+3\frac{\kappa_2^2}{\kappa_1^4}-\frac{6\kappa_3}{\kappa_1^3}\right)$$
$$+\frac{1}{n(n-1)}\left(\frac{3}{4}\frac{\kappa_4}{\kappa_2^2}-\frac{3\kappa_3}{\kappa_2\kappa_1}+\frac{3\kappa_2}{\kappa_1^2}\right)+\frac{1}{(n-1)^2}\left(\frac{3}{4}\right).$$

For samples from a symmetrical population:

$$\mu_1'\left(\frac{v}{V}\right) = 1+\frac{1}{n}\left(-\frac{1}{8}\frac{\kappa_4}{\kappa_2^2}+\frac{\kappa_2}{\kappa_1^2}\right)-\frac{1}{4(n-1)}+\frac{1}{n^2}\left(-\frac{15}{128}\frac{\kappa_4^2}{\kappa_2^4}+\frac{\kappa_6}{16\kappa_2^3}+\frac{3}{8}\frac{\kappa_4}{\kappa_1^2\kappa_2}+\frac{3\kappa_2^2}{\kappa_1^4}\right)$$
$$+\frac{1}{n(n-1)}\left(\frac{9}{32}\frac{\kappa_4}{\kappa_2^2}-\frac{1}{4}\frac{\kappa_2}{\kappa_1^2}\right)+\frac{1}{32(n-1)^2},$$

$$\mu_2\left(\frac{v}{V}\right) = \frac{1}{n}\left(\frac{1}{4}\frac{\kappa_4}{\kappa_2^2}+\frac{\kappa_2}{\kappa_1^2}\right)+\frac{1}{2(n-1)}+\frac{1}{n^2}\left(\frac{7}{32}\frac{\kappa_4^2}{\kappa_2^4}-\frac{\kappa_6}{8\kappa_2^3}+\frac{5}{2}\frac{\kappa_4}{\kappa_1^2\kappa_2}+\frac{8\kappa_2^2}{\kappa_1^4}\right)-\frac{1}{n(n-1)}\left(\frac{5}{8}\frac{\kappa_4}{\kappa_2^2}-\frac{\kappa_2}{\kappa_1^2}\right)-\frac{1}{8(n-1)^2},$$

$$\mu_3\left(\frac{v}{V}\right) = \frac{1}{n^2}\left(-\frac{3}{16}\frac{\kappa_4^2}{\kappa_2^4}+\frac{\kappa_6}{8\kappa_2^3}+\frac{3\kappa_4}{\kappa_1^2\kappa_2}+\frac{6\kappa_2^2}{\kappa_1^4}\right)+\frac{1}{n(n-1)}\left(\frac{3}{4}\frac{\kappa_4}{\kappa_2^2}+\frac{3\kappa_2}{\kappa_1^2}\right)+\frac{1}{4(n-1)^2},$$

$$\mu_4\left(\frac{v}{V}\right) = \frac{1}{n^2}\left(\frac{3}{16}\frac{\kappa_4^2}{\kappa_2^4}+\frac{3}{2}\frac{\kappa_4}{\kappa_1^2\kappa_2}+\frac{3\kappa_2^2}{\kappa_1^4}\right)+\frac{1}{n(n-1)}\left(\frac{3}{4}\frac{\kappa_4}{\kappa_2^2}+\frac{3\kappa_2}{\kappa_1^2}\right)+\frac{3}{4(n-1)^2}.$$

For samples from a normal population:

$$\mu_1'\left(\frac{v}{V}\right) = 1+\frac{V^2}{n}-\frac{1}{4(n-1)}+\frac{3V^4}{n^2}-\frac{V^2}{4n(n-1)}+\frac{1}{32(n-1)^2},$$

$$\mu_2\left(\frac{v}{V}\right) = \frac{V^2}{n}+\frac{1}{2(n-1)}+\frac{8V^4}{n^2}-\frac{1}{8(n-1)^2}+\frac{V^2}{n(n-1)},$$

$$\mu_3\left(\frac{v}{V}\right) = \frac{6V^4}{n^2}+\frac{3V^2}{n(n-1)}+\frac{1}{4(n-1)^2},$$

$$\mu_4\left(\frac{v}{V}\right) = \frac{3V^4}{n^2}+\frac{3V^2}{n(n-1)}+\frac{3}{4(n-1)^2}.$$

10. In order that the truncation of the parent population at the point $x = 0$ should play as small a part as possible, it will be necessary to choose V small. Thus for the case of a normal parent population it is assumed that V can never be greater than $\frac{1}{3}$, or in other words it is assumed that the truncation is made at a point farther away from the mean than -3σ. It is understood that in practical problems it is rare to find a V as large as $\frac{1}{3}$. The manipulative algebra involved in obtaining the approximations to the moments given above is elementary but heavy, and it would be satisfactory if it were possible to apply a check of some kind. This is difficult for the general case. We may note that for the normal population, with V small and n large, we have

$$\mu_2\left(\frac{v}{V}\right) \doteq \frac{1}{2n}(1+2V^2),$$

a value given, it is believed, originally by K. Pearson. For the general case Kendall (1943) gives

$$\mu_2\left(\frac{v}{V}\right) = \frac{1}{n}\left(\frac{\mu_4-\mu_2^2}{4\mu_2^2}+\frac{\mu_2}{\mu_1'^2}-\frac{\mu_3}{\mu_2\mu_1'}\right) = \frac{1}{n}\left[\frac{\kappa_4}{4\kappa_2^2}+\frac{\kappa_2}{\kappa_1^2}-\frac{\kappa_3}{\kappa_2\kappa_1}\right]+\frac{1}{2n},$$

which, for large n, agrees with the leading terms of the expression for $\mu_2(v/V)$ for the general case given in the preceding section. These small checks were all that could be found.

11. In spite of the restrictions made necessary in order to use the expansion, it is possible to get some idea of the distribution of v/V from the approximate moments, particularly if all that is required is to gauge the approximate sample size for which it is reasonable to assume that v is normally distributed. We take the case where the parent population is normal; the momental constants for selected values of v/V and n are given in Table 2.

Table 2. *Momental constants of the distribution of v/V when the parent population is normal*

Sample size	$V=\frac{1}{3}$				$V=\frac{1}{5}$				$V=\frac{1}{10}$			
n	μ_1'	σ	β_1	β_2	μ_1'	σ	β_1	β_2	μ_1'	σ	β_1	β_2
6	0·970	0·346	0·31	2·93	0·956	0·321	0·18	3·20	0·951	0·311	0·13	3·29
11	0·9855	0·246	0·17	2·94	0·9786	0·230	0·09	3·09	0·9759	0·223	0·06	3·14
21	0·9929	0·175	0·09	2·97	0·9894	0·163	0·05	3·04	0·9880	0·159	0·03	3·07
31	0·9953	0·143	0·06	2·98	0·9930	0·134	0·03	3·03	0·9920	0·130	0·02	3·05
41	0·9965	0·124	0·05	2·98	0·9947	0·116	0·02	3·02	0·9940	0·113	0·02	3·03

It is clear that the distribution of v/V as judged from the approximate moments tends to normality reasonably quickly as the sample size is increased, and that for samples of over 40 little error will be made in assuming that v/V is normally distributed.

12. McKay (1932) considered the distribution of

$$n\left(1+\frac{1}{V^2}\right)\left(\frac{\hat{r}^2}{1+\hat{r}^2}\right),$$

where

$$\hat{r} = \frac{s}{\bar{x}} \quad \text{and} \quad s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2.$$

He showed that this quantity is approximately distributed as χ^2 with $n-1$ degrees of freedom when the parent population is normal. In the previous sections of this paper we have been considering

$$v = \frac{s_c}{\bar{x}},$$

where s_c^2 is the unbiased estimate of the population variance. It follows therefore that we might expect

$$(n-1)\left(1+\frac{1}{V^2}\right)\left(\frac{v^2}{1+v^2}\right)$$

to be approximately distributed as χ^2 with $n-1$ degrees of freedom in samples from a normal population. We may use this approximation of McKay to check the adequacy of the assumption that the distribution of v/V is normal when that of the parent population generating the sample is normal. The procedure is as follows. We choose a significance level for χ^2, say 5 % for the sake of illustration, and for a given n and V we calculate $v_{0\cdot05}$ corresponding to this level. Hence, assuming v/V is normally distributed, we may calculate the tail area corresponding to this $v_{0\cdot05}$. The results of such calculations for two values of n and three values of V are given in Table 3. They again suggest that the assumption of normality for v/V is not likely to lead to serious error in tests of significance, provided the sample is of reasonable size.

Table 3. *Estimated probability integral of the distribution of v/V corresponding to $\chi^2_{0\cdot05}$*

n	$V=\frac{1}{3}$	$V=\frac{1}{5}$	$V=\frac{1}{10}$
101	0·047	0·048	0·049
41	0·045⁻	0·046	0·047

13. It is seen (for example in Table 2) that as n increases, the distribution of v/V becomes less and less dependent on V. We therefore look at the possibility of abbreviating the expression for the second moment. Four possibilities may be considered. We assume that for a normal parent population v/V is normally distributed about unity with standard errors equal to

$$\text{(i)}\ \sqrt{\frac{1}{2n}},\quad \text{(ii)}\ \frac{1}{\sqrt{(2n)}}(1+2V^2)^{\frac{1}{2}},\quad \text{(iii)}\ \left(\frac{V^2}{n}+\frac{1}{2(n-1)}\right)^{\frac{1}{2}},$$

$$\text{(iv)}\ \left(\frac{V^2}{n}+\frac{1}{2(n-1)}+\frac{8V^4}{n^2}-\frac{1}{8(n-1)^2}+\frac{V^2}{n(n-1)}\right)^{\frac{1}{2}}.$$

The results for sample size 41 and different values of V are given in Table 4.

Table 4. *Various estimates of standard error of v/V for samples of size* 41

Expression for standard error	$V=\frac{1}{3}$	$V=\frac{1}{5}$	$V=\frac{1}{10}$
(i)	0·110	0·110	0·110
(ii)	0·122	0·115	0·112
(iii)	0·123	0·116	0·113
(iv)	0·124	0·116	0·113

Approximation (i) is clearly inadequate for the size of sample considered, (ii) is not far off (iii) and (iv), and will probably be good enough for rough tests of significance. Actually, since it gives values less than the values obtained from (iv), the probability integral calculated from it will give values closer to the true significance level as judged by the χ^2 distribution.

14. It will be seen, as might indeed be expected, that there is a similarity between the moments of the distribution of the standard deviation estimated from normal samples and those of the distribution of v/V. Approximations to the moments of the distribution of the standard deviation can be found quickly, using the k-statistic technique. Writing

$$s_e^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2,$$

then
$$\mathscr{E}\left(\frac{s_e}{\kappa_2^{\frac{1}{2}}}\right) = \mathscr{E}\left(1 + \frac{k_2 - \kappa_2}{\kappa_2}\right)^{\frac{1}{2}} = 1 + \frac{1}{2\kappa_2}(\overline{2}) - \frac{1}{8\kappa_2^2}(\overline{2}^2) + \frac{1}{16\kappa_2^3}(\overline{2}^3) - \frac{5}{128\kappa_2^4}(\overline{2}^4) + \ldots$$

$$= 1 - \frac{8\kappa_4}{n\kappa_2^2} - \frac{1}{4(n-1)} + \frac{1}{n^2}\left(\frac{\kappa_6}{16\kappa_2^3} - \frac{15}{128}\frac{\kappa_4^2}{\kappa_2^4}\right) + \frac{1}{n(n-1)}\left(\frac{9}{32}\frac{\kappa_4}{\kappa_2^2}\right) + \frac{1}{(n-1)^2}\left(\frac{1}{32} + \frac{\kappa_3^2}{4\kappa_2^3}\right),$$

and for a normal population we have

$$\mathscr{E}\left(\frac{s_e}{\kappa_2^{\frac{1}{2}}}\right) = 1 - \frac{1}{4(n-1)} + \frac{1}{32(n-1)^2}, \quad \mu_2\left(\frac{s_e}{\kappa_2^{\frac{1}{2}}}\right) = \frac{1}{2(n-1)} - \frac{1}{8(n-1)^2},$$

$$\mu_3\left(\frac{s_e}{\kappa_2^{\frac{1}{2}}}\right) = \frac{1}{4(n-1)^2}, \quad \mu_4\left(\frac{s_e}{\kappa_2^{\frac{1}{2}}}\right) = \frac{3}{4(n-1)^2}.$$

Thus the moments of $s_e/\kappa_2^{\frac{1}{2}}$ are those of v/V, in which V is put equal to zero in the final expansion. Since the distribution of s_e tends fairly quickly to normality it may be expected that the distribution of v/V will do likewise.

15. We return to the approximate moments of v/V and consider the case when the parent population is not normal. It will be assumed that the parent population may be described by a Gram-Charlier Series, Type A, and we write for its functional form

$$y = \frac{1}{\kappa_2^{\frac{1}{2}}\sqrt{(2\pi)}} \exp\left[-\frac{1}{2}\left(\frac{X-\kappa_1}{\kappa_2^{\frac{1}{2}}}\right)^2\right]\left(1 + \frac{\sqrt{\beta_1}}{3!}H_3\left(\frac{X-\kappa_1}{\kappa_2^{\frac{1}{2}}}\right) + \frac{\beta_2}{4!}H_4\left(\frac{X-\kappa_1}{\kappa_2^{\frac{1}{2}}}\right)\right),$$

where $H_3(x)$ and $H_4(x)$ are the Hermite polynomials of orders 3 and 4 respectively. This population has cumulants

$$\kappa_1,\ \kappa_2,\ \kappa_3,\ \kappa_4,\ \kappa_5 = 0, \quad \kappa_6 = -10\kappa_3^2, \quad \kappa_7 = -35\kappa_4\kappa_3, \quad \kappa_8 = -35\kappa_4^2, \quad \text{and so on.}$$

If we put $\kappa_4 = 0$ in this population, we may eliminate any effect due to kurtosis, and similarly if we give κ_4 a non-zero value and let $\kappa_3 = 0$ we can study the effect of kurtosis when there is no asymmetry. The moments of the criterion v/V may be written in simple forms for these two cases. Writing the moments of v/V for the normal population, that is, when $\kappa_3 = \kappa_4 = 0$, as $\mu_1'(v/V \mid N)$, $\mu_2(v/V \mid N)$, ..., and putting

$$\beta_1 = \frac{\kappa_3^2}{\kappa_2^3}, \quad \beta_2 - 3 = \frac{\kappa_4}{\kappa_2^2} = \gamma_2,$$

we have the following expressions:

Moments of v/V when $\kappa_4 = 0$ in the Gram-Charlier population:

$$\mu_1'\left(\frac{v}{V}\right) = \mu_1'\left(\frac{v}{V}\middle| N\right) + \beta_1\left(\frac{1}{4(n-1)^2} - \frac{5}{8n^2} - \frac{V^2}{4n^2}\right) + V\sqrt{\beta_1}\left(-\frac{1}{2n} + \frac{1}{8n(n-1)} - \frac{5V^2}{2n^2}\right),$$

$$\mu_2\left(\frac{v}{V}\right) = \mu_2\left(\frac{v}{V}\middle| N\right) + \beta_1\left(\frac{5}{4n^2} - \frac{1}{2(n-1)^2} + \frac{V^2}{4n^2}\right) - V\sqrt{\beta_1}\left(\frac{1}{n} + \frac{1}{2n(n-1)} + \frac{10V^2}{n^2}\right),$$

$$\mu_3\left(\frac{v}{V}\right) = \mu_3\left(\frac{v}{V}\middle| N\right) + \beta_1\left(-\frac{5}{4n^2} + \frac{1}{2(n-1)^2} + \frac{9V^2}{n^2}\right) - V\sqrt{\beta_1}\left(\frac{3}{n(n-1)} + \frac{10V^2}{n^2}\right),$$

$$\mu_4\left(\frac{v}{V}\right) = \mu_4\left(\frac{v}{V}\middle| N\right) + \beta_1\left(\frac{3V^2}{n^2}\right) - V\sqrt{\beta_1}\left(\frac{3}{n(n-1)} + \frac{6V^2}{n^2}\right).$$

Moments of v/V when $\kappa_3 = 0$ in the Gram-Charlier population:

$$\mu_1'\left(\frac{v}{V}\right) = \mu_1'\left(\frac{v}{V}\middle| N\right) + \gamma_2\left[-\frac{1}{8n} + \frac{9}{32n(n-1)} + \frac{3V^2}{8n^2}\right] - \frac{15}{128n^2}\gamma_2^2,$$

$$\mu_2\left(\frac{v}{V}\right) = \mu_2\left(\frac{v}{V}\middle| N\right) + \gamma_2\left[\frac{1}{4n} - \frac{5}{8n(n-1)} + \frac{5V^2}{2n^2}\right] + \frac{7}{32n^2}\gamma_2^2,$$

$$\mu_3\left(\frac{v}{V}\right) = \mu_3\left(\frac{v}{V}\middle| N\right) + \gamma_2\left[\frac{3}{4n(n-1)} + \frac{3V^2}{n^2}\right] - \frac{3}{16n^2}\gamma_2^2,$$

$$\mu_4\left(\frac{v}{V}\right) = \mu_4\left(\frac{v}{V}\middle| N\right) + \gamma_2\left[\frac{3}{4n(n-1)} + \frac{3V^2}{2n^2}\right] + \frac{3}{16n^2}\gamma_2^2.$$

A third alternative would be to let both κ_3 and κ_4 have given values, and to study the effect of combined asymmetry and kurtosis in the parent population.

16. The moments of v/V for some parent population which is different from normal having been obtained, the next step is a matter for choice. One may either, from a study of the β_1 and β_2 of the distribution of v/V fit the appropriate Pearson curve using the first four moments, or one may choose for v/V a distribution having the same functional form as the parent population. In either case the procedure enables an estimate of the probability integral to be made.

17. Pearl (1905) gives the momental constants for the distribution of brain-weights in a number of races. We take the first line of his table and note that 413 Swedish males had brains of average weight 1400·481 g. The coefficient of variation was 0·07592, and the β_1 and β_2 of the distribution 0·0287 and 2·7964 respectively. We assume that these momental constants are those of the true distribution of Swedish male brain-weights. We then ask the following question: if a sample of 51 observations is available and known to come from the population with $V = 0{\cdot}07592$, what difference does it make to the expected mean and standard error of v/V if we assume the distribution of the population to be normal, instead of, with β_1 and β_2 as given? Substituting in the general formula we have

$$\mu_1'(v/V \mid N) = 0{\cdot}9951, \quad \mu_2(v/V \mid N) = 0{\cdot}0101, \quad \sigma_N = 0{\cdot}100;$$
$$\mu_1'(v/V \mid \beta_1\beta_2) = 0{\cdot}9955, \quad \mu_2(v/V \mid \beta_1\beta_2) = 0{\cdot}0089, \quad \sigma_{\beta_1\beta_2} = 0{\cdot}094.$$

For this population, therefore, the mean and standard deviation of v/V appear to be little changed from those which would have been obtained if normality had been assumed. It appears likely that in the case of anthropometric data, where the deviations from normality are never very marked and the sample size is usually fairly large, that the expressions for the mean and standard deviation of v/V assuming the parent population is normal will be adequate.

I should like to thank Professor E. S. Pearson and Professor M. G. Kendall for helpful criticism.

Table 1. *Terms to be added to* $\kappa(r^h s^l)$ *to give* $(\overline{r^h s^l})$

$h+l=4$

$\kappa(r^h s^l)$	Add.
$\kappa(1^4)$	$\frac{3\kappa_2^2}{n^2}$
$\kappa(2\,1^3)$	$\frac{3\kappa_2\kappa_3}{n^2}$
$\kappa(2^2\,1^2)$	$\frac{\kappa_4\kappa_2}{n^2}+\frac{2\kappa_3^2}{n^2}+\frac{2\kappa_2^3}{n(n-1)}$
$\kappa(2^3\,1)$	$\frac{3\kappa_4\kappa_2}{n^2}+\frac{6\kappa_3\kappa_2^2}{n(n-1)}$
$\kappa(2^4)$	$\frac{3\kappa_4^2}{n^2}+\frac{12\kappa_4\kappa_2^2}{n(n-1)}+\frac{12\kappa_2^4}{(n-1)^2}$
$\kappa(3\,1^3)$	$\frac{3\kappa_4\kappa_2}{n^2}$
$\kappa(3^2\,1^2)$	$\frac{\kappa_6\kappa_2}{n^2}+\frac{2\kappa_4^2}{n^2}+\frac{9\kappa_4\kappa_2^2}{n(n-1)}+\frac{9\kappa_3^2\kappa_2}{n(n-1)}+\frac{6\kappa_2^4}{(n-1)(n-2)}$
$\kappa(3^3\,1)$	$\frac{3\kappa_6\kappa_4}{n^2}+\frac{27\kappa_4^2\kappa_2}{n(n-1)}+\frac{27\kappa_4\kappa_3^2}{n(n-1)}+\frac{18\kappa_4\kappa_2^3}{(n-1)(n-2)}$
$\kappa(3^4)$	$\frac{3\kappa_6^2}{n^2}+\frac{54\kappa_6\kappa_4\kappa_2}{n(n-1)}+\frac{54\kappa_6\kappa_3^2}{n(n-1)}+\frac{36\kappa_6\kappa_2^3}{(n-1)(n-2)}+\frac{486\kappa_4\kappa_3^2\kappa_2}{(n-1)^2}+\frac{324\kappa_4\kappa_2^4 n}{(n-1)^2(n-2)}+\frac{324\kappa_3^2\kappa_2^3 n}{(n-1)^2(n-2)}+\frac{243\kappa_4^2\kappa_2^2}{(n-1)^2}+\frac{243\kappa_3^2\kappa_2^2}{(n-1)^2}+\frac{108\kappa_2^6 n^2}{(n-1)^3(n-2)^2}$
$\kappa(4\,1^3)$	$\frac{3\kappa_5\kappa_2}{n^2}$
$\kappa(4^2\,1^2)$	$\frac{\kappa_8\kappa_2}{n^2}+\frac{16\kappa_6\kappa_2^2}{n(n-1)}+\frac{2\kappa_5^2}{n^2}+\frac{48\kappa_5\kappa_3\kappa_2}{n(n-1)}+\frac{34\kappa_4^2\kappa_2}{n(n-1)}+\frac{72\kappa_4\kappa_2^3}{(n-1)(n-2)}+\frac{144\kappa_3^2\kappa_2^2}{(n-1)(n-2)}+\frac{24\kappa_2^5(n+1)}{(n-1)(n-2)(n-3)}$
$\kappa(4^3\,1)$	$\frac{3\kappa_8\kappa_2}{n^2}+\frac{48\kappa_6\kappa_5\kappa_2}{n(n-1)}+\frac{144\kappa_5^2\kappa_3}{n(n-1)}+\frac{102\kappa_5\kappa_4^2}{n(n-1)}+\frac{216\kappa_5\kappa_4\kappa_2^2}{(n-1)(n-2)}+\frac{432\kappa_5\kappa_3^2\kappa_2}{(n-1)(n-2)}+\frac{72\kappa_5\kappa_2^4(n+1)}{(n-1)(n-2)(n-3)}$
$\kappa(3\,2^3)$	$\frac{3\kappa_6\kappa_4}{n^2}+\frac{6\kappa_6\kappa_2^2}{n(n-1)}+\frac{18\kappa_4\kappa_3\kappa_2}{n(n-1)}+\frac{36}{(n-1)^2}\kappa_3\kappa_2^2$
$\kappa(3^2\,2^2)$	$\frac{\kappa_8\kappa_4}{n^2}+\frac{2\kappa_6\kappa_2^2}{n(n-1)}+\frac{2\kappa_5^2}{n^2}+\frac{24\kappa_5\kappa_3\kappa_2}{n(n-1)}+\frac{9\kappa_4^2\kappa_2}{n(n-1)}+\frac{9\kappa_4\kappa_3^2}{n(n-1)}+\frac{18\kappa_4\kappa_2^3}{(n-1)^2}+\frac{6\kappa_4\kappa_2^3}{(n-1)(n-2)}+\frac{90\kappa_3^2\kappa_2^2}{(n-1)^2}+\frac{12\kappa_2^5 n}{(n-1)^2(n-2)}$
$\kappa(3^3\,2)$	$\frac{3\kappa_8\kappa_5}{n^2}+\frac{18\kappa_6\kappa_3\kappa_2}{n(n-1)}+\frac{27\kappa_5\kappa_4\kappa_2}{n(n-1)}+\frac{162\kappa_4\kappa_3\kappa_2^2}{(n-1)^2}+\frac{27\kappa_6\kappa_3^2}{n(n-1)}+\frac{18\kappa_5\kappa_2^3}{(n-1)(n-2)}+\frac{162\kappa_3^2\kappa_2}{(n-1)^2}+\frac{108n\kappa_3\kappa_2^4}{(n-1)^2(n-2)}$

$h+l=5$

$\kappa(r^h s^l)$	Add.
$\kappa(1^5)$	$\frac{10\kappa_3\kappa_2}{n^3}$
$\kappa(2\,1^4)$	$\frac{6\kappa_4\kappa_2}{n^3}+\frac{4\kappa_3^2}{n^3}$
$\kappa(2^2\,1^3)$	$\frac{3\kappa_6\kappa_2}{n^3}+\frac{7\kappa_4\kappa_3}{n^3}+\frac{14\kappa_3\kappa_2^2}{n^2(n-1)}$
$\kappa(2^3\,1^2)$	$\frac{\kappa_8\kappa_2}{n^3}+\frac{6\kappa_5\kappa_3}{n^3}+\frac{3\kappa_4^2}{n^3}+\frac{18\kappa_4\kappa_2^2}{n^2(n-1)}+\frac{(28n-32)\kappa_3^2\kappa_2}{n^2(n-1)^2}+\frac{8}{n(n-1)^2}\kappa_2^4$
$\kappa(2^4\,1)$	$\frac{4\kappa_8\kappa_2}{n^3}+\frac{6\kappa_5\kappa_4}{n^3}+\frac{12\kappa_5\kappa_2^2}{n^2(n-1)}+\frac{72\kappa_4\kappa_3\kappa_2}{n^2(n-1)}+\frac{16(n-2)\kappa_3^3}{n^2(n-1)^2}+\frac{80\kappa_3\kappa_2^3}{n(n-1)^2}$
$\kappa(2^5)$	$\frac{10\kappa_8\kappa_4}{n^3}+\frac{20\kappa_6\kappa_2^2}{n^2(n-1)}+\frac{120\kappa_4^2\kappa_2}{n^2(n-1)}+\frac{320\kappa_4\kappa_3^2}{n(n-1)^2}+\frac{40(n-2)}{n^2(n-1)^2}\kappa_4\kappa_2^3+\frac{80(n-2)}{n(n-1)^3}\kappa_3^2\kappa_2^2+\frac{160\kappa_2^5}{(n-1)^3}$

Table 1. (*cont.*)

$h+l=5$

$\kappa(3\,1^4)$ $\dfrac{6\kappa_5\kappa_2}{n^3}+\dfrac{4\kappa_3\kappa_4}{n^3}$

$\kappa(3^2\,1^3)$ $\dfrac{3\kappa_7\kappa_2}{n^3}+\dfrac{\kappa_6\kappa_3}{n^3}+\dfrac{6\kappa_5\kappa_4}{n^3}+\dfrac{27\kappa_5\kappa_3^2}{n^2(n-1)}+\dfrac{90\kappa_4\kappa_3\kappa_2}{n^2(n-1)}+\dfrac{9\kappa_3^3}{n^2(n-1)}+\dfrac{60\kappa_3\kappa_2^2}{n(n-1)(n-2)}$

$\kappa(3^3\,1^2)$ $\dfrac{6\kappa_7\kappa_4}{n^3}+\dfrac{\kappa_6\kappa_2}{n^3}+\dfrac{27}{n^2(n-1)}\kappa_7\kappa_2^2+\dfrac{3\kappa_5\kappa_6}{n^3}+\dfrac{27(3n-4)\kappa_6\kappa_3\kappa_2}{n^2(n-1)^2}+\kappa_5\kappa_4\kappa_2\left(\dfrac{81}{n^2(n-1)}+\dfrac{27(4n-7)}{n^2(n-1)^2}\right)$
$+\dfrac{27}{n^2(n-1)}\kappa_5\kappa_3^2+\kappa_5\kappa_2^2\left(\dfrac{18}{n(n-1)(n-2)}+\dfrac{54(4n-7)}{n(n-1)^2(n-2)}\right)+\dfrac{162\kappa_4^2\kappa_3}{n^2(n-1)}$
$+\kappa_4\kappa_3\kappa_2^2\left(\dfrac{165(5n-12)}{n(n-1)^2(n-2)}+\dfrac{108}{n(n-1)n(-2)}\right)+\dfrac{36(7n^2-30n+34)}{n(n-1)^2(n-2)^2}\kappa_3^2\kappa_2+\dfrac{108(5n-12)}{(n-1)^2(n-2)^2}\kappa_3\kappa_2^4$

$h+l=6$. Corrective terms involving $\dfrac{1}{n^3}$ only

$\kappa(r^s\,h^l)$ Add.

$\kappa(1^6)$ $\dfrac{15\kappa_2^3}{n^3}$

$\kappa(2\,1^5)$ $\dfrac{15\kappa_3\kappa_2^2}{n^3}$

$\kappa(2^2\,1^4)$ $\dfrac{12\kappa_3^2\kappa_2}{n^3}+\dfrac{3\kappa_4\kappa_2^2}{n^3}+\dfrac{6\kappa_2^4}{n^2(n-1)}$

$\kappa(2^3\,1^3)$ $\dfrac{9\kappa_4\kappa_3\kappa_2}{n^3}+\dfrac{6\kappa_3^3}{n^3}+\dfrac{18\kappa_3\kappa_2^3}{n^2(n-1)}$

$\kappa(2^4\,1^2)$ $\dfrac{12\kappa_4\kappa_3^2}{n^3}+\dfrac{24\kappa_3^2\kappa_2^2}{n^2(n-1)}+\dfrac{3\kappa_4^2\kappa_2}{n^3}+\dfrac{12\kappa_4\kappa_2^3}{n^2(n-1)}+\dfrac{12\kappa_2^5}{n(n-1)^2}$

$\kappa(2^5\,1)$ $\dfrac{15\kappa_4^2\kappa_3}{n^3}+\dfrac{60\kappa_4\kappa_3\kappa_2^2}{n^2(n-1)}+\dfrac{60\kappa_3\kappa_2^4}{n(n-1)^2}$

$\kappa(2^6)$ $\dfrac{15\kappa_4^3}{n^3}+\dfrac{90\kappa_4^2\kappa_2^2}{n^2(n-1)}+\dfrac{180\kappa_4\kappa_2^4}{n(n-1)^2}+\dfrac{120\kappa_2^6}{(n-1)^3}$

$\kappa(3\,1^5)$ $\dfrac{15\kappa_4\kappa_2^2}{n^3}$

$\kappa(3^2\,1^4)$ $\dfrac{3\kappa_4^2\kappa_2}{n^3}+\dfrac{12\kappa_4^2\kappa_2}{n^3}+\dfrac{27\kappa_4\kappa_2^3}{n^2(n-1)}+\dfrac{27\kappa_3^2\kappa_2^2}{n^2(n-1)}+\dfrac{18\kappa_2^5}{n(n-1)(n-2)}$

REFERENCES

FISHER, R. A. (1928). *Proc. Lond. Math. Soc.* Ser. 2, **30**, 199.
KENDALL, M. G. (1940*a*). *Ann. Eugen., Lond.*, **10**, 106.
KENDALL, M. G. (1940*b*). *Ann. Eugen., Lond.*, **10**, 215.
KENDALL, M. G. (1940*c*). *Ann. Eugen., Lond.*, **10**, 392.
KENDALL, M. G. (1942). *Ann. Eugen., Lond.*, **11**, 300.
KENDALL, M. G. (1943). *The Advanced Theory of Statistics*. I. Chas. Griffin and Co.
MCKAY, A. T. (1932). *J.R. Statist. Soc.* **96**. 695.
PEARL, R. (1905). *Biometrika*, **4**, 38.
WISHART, J. (1929*a*). *Proc. Lond. Math. Soc.* **29**, 309.
WISHART, J. (1929*b*). *Proc. Roy. Soc. Edinb.* **49**, 78.
WISHART, J. (1930). *Biometrika*, **22**, 224.
WISHART, J. (1933). *Biometrika*, **25**, 52.

THE MOMENTS OF THE z AND F DISTRIBUTIONS

By F. N. DAVID

1. The cumulants of Fisher's z distribution were derived approximately by Cornish & Fisher (1937) and exactly by J. Wishart (1947*a*) some years later. In both cases, however, the assumption is that the parent population (or populations) generating the samples is normal. It appears simple, using Fisher's k-statistic technique, to derive approximations to the cumulants of z and F when the two estimates of variance involved are based on independent samplings from two parent populations which may have any distributions whatever provided the cumulants exist. These approximations to the true cumulants can be used to investigate the effect of non-normality in the parent population on the distributions of z and F, or to obtain approximations to the power of the z and F tests with respect to a set of specifically defined alternative hypotheses. It should be noted that in the type of problem we have in mind, the two estimates of variance are essentially independent or at least uncorrelated. In many problems in the analysis of variance non-normality in the parent population introduces a correlation between the estimates which are independent in the normal case (see E. S. Pearson (1931) and R. C. Geary (1947)). We are not concerned here with this latter problem.

2. It will be assumed that there are two populations π_1 and π_2, each of which may be described by cumulants up to any order desired. For the cumulants of π_1 we write $\kappa_1, \kappa_2, \ldots$, and for those of $\pi_2, \kappa_1', \kappa_2', \ldots$. Following MacMahon we define the one-part symmetric functions as

$$s_r = \sum_{j=1}^{n} x_j^r,$$

where n is the number of magnitudes involved. Hence if we imagine samples of sizes n_1 and n_2 randomly and independently drawn from π_1 and π_2 respectively, and if we associate with each element of the sample a random variable $x_1, x_2, \ldots, x_{n_1}, x_1', x_2', \ldots, x_{n_2}'$, we shall have

$$z = \tfrac{1}{2} \log_e \frac{k_2}{k_2'},$$

where

$$k_2 = \frac{n_1}{n_1 - 1} m_2 = \frac{n_1}{n_1 - 1} \left(\frac{s_2}{n_1} - \frac{s_1^2}{n_1^2} \right)$$

for π_1 and a similar interpretation may be given for k_2' and π_2.

3. We begin by expanding z:

$$2z - \log_e \frac{\kappa_2}{\kappa_2'} = \log_e \left(1 + \frac{k_2 - \kappa_2}{\kappa_2} \right) - \log_e \left(1 + \frac{k_2' - \kappa_2'}{\kappa_2'} \right).$$

It is clear that expansion of the right-hand side will only be valid provided

$$k_2 < 2\kappa_2 \quad \text{and} \quad k_2' < 2\kappa_2'.$$

The general question of the validity of expansions of this type has been attempted only by J. B. D. Derksen (1939), but it appears possible to justify the use of such expansions

for reasonably large n in an approximate sort of way. We argue for the expansion of $\log_e(1+(k_2-\kappa_2)/\kappa_2)$ only, but it is obvious that the same kind of reasoning may also be applied to $\log_e(1+(k_2'-\kappa_2')/\kappa_2')$. k_2 is distributed with mean κ_2 and variance $\sigma^2_{k_2}$, where

$$\sigma^2_{k_2} = \frac{\kappa_4}{n_1} + \frac{2\kappa_2^2}{n_1-1}.$$

For any reasonable sized n_1 it is clear that

$$\left(\frac{\kappa_4}{n_1} + \frac{2\kappa_2^2}{n_1-1}\right)^{\frac{1}{2}} < \kappa_2,$$

and, in fact, n_1 may be so chosen that for some fixed positive integer r,

$$r\left(\frac{\kappa_4}{n_1} + \frac{2\kappa_2^2}{n_1-1}\right)^{\frac{1}{2}} < \kappa_2.$$

Provided $r > 3$ it is clear that for reasonable-sized n_1 we shall have

$$P\{k_2 > \kappa_2 + r\sigma_{k_2}\} < \epsilon,$$

where ϵ is some small positive fraction. Hence it appears, for n_1 of reasonable size, the expansion will be valid, except in a small proportion of cases, for most distributions met with in statistical practice. An alternative approach would be to regard the distribution of k_2 as truncated at the point $k_2 = 2\kappa_2$, the moments derived being applicable to this truncated distribution.

4. If we write

$$2z - \log_e \frac{\kappa_2}{\kappa_2'} = \alpha - \alpha',$$

where

$$\alpha = \log_e\left(1 + \frac{k_2-\kappa_2}{\kappa_2}\right) \quad \text{and} \quad \alpha' = \log_e\left(1 + \frac{k_2'-\kappa_2'}{\kappa_2'}\right),$$

then since α and α' are independent we shall have

$$\mathscr{E}(z) = \tfrac{1}{2}\log_e \frac{\kappa_2}{\kappa_2'} + \tfrac{1}{2}\mathscr{E}(\alpha) - \tfrac{1}{2}\mathscr{E}(\alpha'),$$

$$\kappa_2(z) = \sigma_z^2 = \tfrac{1}{4}(\kappa_2(\alpha) + \kappa_2(\alpha')),$$

$$\kappa_3(z) = \tfrac{1}{8}(\kappa_3(\alpha) - \kappa_3(\alpha')),$$

$$\kappa_4(z) = \tfrac{1}{16}(\kappa_4(\alpha) + \kappa_4(\alpha')).$$

Further, since α and α' have the same functional form, we may derive the moments of α and obtain the moments of α' by adding the appropriate dashes. We consider then

$$\alpha = \log_e\left(1 + \frac{k_2-\kappa_2}{\kappa_2}\right) = \frac{k_2-\kappa_2}{\kappa_2} - \frac{1}{2}\frac{(k_2-\kappa_2)^2}{\kappa_2^2} + \frac{1}{3}\frac{(k_2-\kappa_2)^3}{\kappa_2^3} - \ldots,$$

and take expectations. Using the notation

$$\mathscr{E}(k_2-\kappa_2)^r = (\overline{2^r}),$$

it is immediate that

$$\mathscr{E}(\alpha) = -\frac{1}{2\kappa_2^2}(\overline{2^2}) + \frac{1}{3\kappa_2^3}(\overline{2^3}) - \frac{1}{4\kappa_2^4}(\overline{2^4}) + \ldots.$$

From Fisher's or Kendall's tables of product cumulants and the tables of corrective factors recently given (David, 1949), we may write down immediately the right-hand side of the above expression, neglecting terms of higher order than $1/n_1^3$:

$$\mathscr{E}(\alpha) = -\frac{1}{n_1-1} - \frac{1}{3(n_1-1)^2} - \frac{\kappa_4}{\kappa_2^2}\left(\frac{n_1^2-4n_1+7}{2n_1(n_1-1)^2}\right) + \frac{\kappa_3^2}{\kappa_2^3}\left(\frac{4(n_1-2)(n_1-7)}{3n_1(n_1-1)^3}\right) + \frac{\kappa_6}{\kappa_2^3}\left(\frac{n_1-7}{3n_1^2(n_1-1)}\right)$$
$$-\frac{\kappa_4^2}{\kappa_2^4}\left(\frac{3n_1^3-13n_1^2+9n_1+9}{4n_1^2(n_1-1)^3}\right) - \frac{\kappa_5\kappa_3}{\kappa_2^4}\left(\frac{8(n_1-2)}{n_1^2(n_1-1)^2}\right) - \frac{\kappa_8}{\kappa_2^4}\frac{1}{4n_1^3} + \frac{\kappa_4\kappa_3^2}{\kappa_2^5}\frac{8(n_1-2)}{n_1^2(n_1-1)^2}$$
$$+\frac{\kappa_6\kappa_4}{\kappa_2^5}\frac{2}{n_1^3} - \frac{\kappa_4^3}{\kappa_2^6}\frac{5}{2n_1^3}.$$

We introduce the notation $\gamma_r = \frac{\kappa_{r+2}}{\kappa_2^{\frac{1}{2}r+1}}$, $f_1 = n_1 - 1$,

and write the expansion of $\mathscr{E}(\alpha)$ in terms of γ and f_1 up to order $1/f_1^3$. Thus

$$\mathscr{E}(\alpha) = -\frac{1}{2f_1}(2+\gamma_2) + \frac{1}{12f_1^2}(-4+18\gamma_2+16\gamma_1^2+4\gamma_4-9\gamma_2^2)$$
$$+\frac{1}{12f_1^3}[-42\gamma_2-128\gamma_1^2-32\gamma_4+30\gamma_2^2-96\gamma_3\gamma_1+96\gamma_2\gamma_1^2-3\gamma_6+24\gamma_4\gamma_2-30\gamma_2^3].$$

The expression for $\mathscr{E}(\alpha')$ will be similar except that f will carry the subscript 2 and the γ's will carry dashes. If we write

$$\left[\frac{\gamma_r}{f^s}\right]_{-2}^{1} = \frac{\gamma_r}{f_1^s} - \frac{\gamma_r'}{f_2^s}$$

we shall have

$$\mathscr{E}(z) = \tfrac{1}{2}\log\frac{\kappa_2}{\kappa_2'} - \left[\frac{1}{4f}(2+\gamma_2)\right]_{-2}^{1} + \frac{1}{24}\left[\frac{1}{f^2}(-4+18\gamma_2+16\gamma_1^2+4\gamma_4-9\gamma_2^2)\right]_{-2}^{1}$$
$$+\frac{1}{24}\left[\frac{1}{f^3}(-42\gamma_2-128\gamma_1^2-32\gamma_4+30\gamma_2^2-96\gamma_3\gamma_1+96\gamma_2\gamma_1^2-3\gamma_6+24\gamma_4\gamma_2-30\gamma_2^3)\right]_{-2}^{1}.$$

5. The higher moments of z follow in a similar way. We use the expansion

$$(\log(1+x))^r = A_{r,0}x^r - A_{r+1,1}\frac{x^{r+1}}{r+1} + A_{r+2,2}\frac{x^{r+2}}{(r+1)(r+2)} - A_{r+3,3}\frac{x^{r+3}}{(r+1)(r+2)(r+3)} + \ldots,$$

where the coefficients $A_{r,s}$ are given by Table 1.

Table 1. *Values of* $A_{r,s}$

s \ r	1	2	3	4	5	6
0	1	1	1	1	1	1
1	.	1	3	6	10	15
2	.	.	2	11	35	85
3	.	.	.	6	50	225

For $\mathscr{E}(\alpha^2)$ we have

$$\mathscr{E}(\alpha^2) = \frac{1}{\kappa_2^2}(\overline{2^2}) - \frac{1}{\kappa_2^3}(\overline{2^3}) + \frac{11}{12\kappa_2^4}(\overline{2^4}) - \frac{5}{6\kappa_2^5}(\overline{2^5}) + \ldots,$$

whence on substitution

$$\mathscr{E}(\alpha^2) = \frac{2}{n_1-1} + \frac{3}{(n_1-1)^2} + \frac{2}{(n_1-1)^3} + \frac{\kappa_4}{\kappa_2^2}\frac{(3n_1^2-9n_1+13)}{3n_1(n_1-1)^2} - \frac{\kappa_6}{\kappa_2^3}\frac{(3n_1-19)}{3n_1^2(n_1-1)} - \frac{\kappa_3^2}{\kappa_2^3}\frac{4(n_1-2)(3n_1-19)}{3n_1(n_1-1)^3}$$
$$+ \frac{\kappa_4^2}{\kappa_2^4}\frac{(33n_1^3-125n_1^2+63n_1+117)}{12n_1^3(n_1-1)^3} + \frac{\kappa_8}{\kappa_2^4}\frac{11}{12n_1^3} - \frac{\kappa_6\kappa_4}{\kappa_2^5}\frac{25}{3n_1^3} + \frac{\kappa_5\kappa_3}{\kappa_2^4}\frac{88(n_1-2)}{3n_1^3(n_1-1)^2}$$
$$- \frac{\kappa_4\kappa_3^2}{\kappa_2^5}\frac{100(n_1-2)}{3n_1^3(n_1-1)^2} + \frac{137}{12n_1^3}\frac{\kappa_4^3}{\kappa_2^6}.$$

We may expand as before in terms of the γ's and powers of f_1, from this write down $\mathscr{E}(\alpha'^2)$ immediately, correct each expansion for the origin not being at the mean, and finally obtain

$$\sigma_z^2 = \frac{1}{4}\left[\frac{1}{f}(2+\gamma_2)\right]_{+2}^{1} + \frac{1}{8}\left[\frac{1}{f^2}(4-6\gamma_2-8\gamma_1^2-2\gamma_4+5\gamma_2^2)\right]_{+2}^{1}$$
$$+ \frac{1}{48}\left[\frac{1}{f^3}(16+84\gamma_2+384\gamma_1^2+96\gamma_4-92\gamma_2^2+352\gamma_3\gamma_1-384\gamma_2\gamma_1^2+11\gamma_6-96\gamma_4\gamma_2+128\gamma_2^3)\right]_{+2}^{1}.$$

Further expansion and algebraic reduction gives

$$\kappa_3(z) = \frac{1}{8}\left[\frac{1}{f^2}(-4+4\gamma_1^2+\gamma_4-3\gamma_2^2)\right]_{-2}^{1}$$
$$+ \frac{1}{16}\left[\frac{1}{f^3}(-16-64\gamma_1^2-16\gamma_4+12\gamma_2^2-96\gamma_3\gamma_1+120\gamma_2\gamma_1^2-3\gamma_6+30\gamma_4\gamma_2-44\gamma_2^3)\right]_{-2}^{1}$$
$$\kappa_4(z) = \frac{1}{16}\left[\frac{1}{f^3}(16+8\gamma_2^2+32\gamma_3\gamma_1-48\gamma_2\gamma_1^2+\gamma_6-12\gamma_4\gamma_2+20\gamma_2^3)\right]_{+2}^{1}.$$

It is clear that to obtain an accurate result for $\kappa_4(z)$ it will be necessary to take further terms in the original expansion in order to be able to obtain the result to order $1/f^4$.

6. The arithmetic involved in the expansions is heavy, and it does not seem possible to obtain a general check. We may note that if the two samples are assumed to have been drawn from different normal populations, then

$$\mathscr{E}(z) - \tfrac{1}{2}\log_e\frac{\kappa_2}{\kappa_2'} = -\frac{1}{2f_1} + \frac{1}{2f_2} - \frac{1}{6f_1^2} + \frac{1}{6f_2^2},$$
$$\kappa_2(z) = \frac{1}{2f_1} + \frac{1}{2f_2} + \frac{1}{2f_1^2} + \frac{1}{2f_2^2} + \frac{1}{3f_1^3} + \frac{1}{3f_2^3},$$
$$\kappa_3(z) = -\frac{1}{2f_1^2} + \frac{1}{2f_2^2} - \frac{1}{f_1^3} + \frac{1}{f_2^3},$$
$$\kappa_4(z) = \frac{1}{f_1^3} + \frac{1}{f_2^3}.$$

When $\kappa_2 = \kappa_2'$, or when the samples may be supposed to have been drawn from the same normal population, the results agree with those given by Cornish & Fisher* (1938). We note

* The Cornish-Fisher results can easily be calculated to any order in $1/f$ that is desired. See Wishart, (1947*a*, p. 174).

the interesting property of the z-distribution, that all cumulants except the first are unaltered by a change in the ratio of the variances of the two normal parent populations, a fact which may be used to determine the power of the variance-ratio test against a set of specifically defined alternative hypotheses. As a further small check we may take only terms of order $1/n$ in σ_z^2, whence

$$\sigma_z^2 = \frac{\beta_2 - 1}{4}\left(\frac{1}{n_1} + \frac{1}{n_2}\right),$$

a result derived by Geary (1947).

7. The exact values of the cumulants for the case when the parent population is normal have been given by Wishart (1947*a*). Further to this he showed that a closer approximation to the true cumulants (after the first) was obtained by expanding in powers of $1/(f-1)$ rather than in powers of $1/f$. For numerical work we shall therefore rewrite the cumulants in powers of $1/(f-1)$. Let

$$r = f - 1 = n - 2.$$

Then

$$\mathscr{E}(z) - \tfrac{1}{2}\log_e \frac{\kappa_2}{\kappa_2'} = -\left[\frac{1}{4r}(2+\gamma_2)\right]_{-2}^{1} + \left[\frac{1}{24r^2}(8+24\gamma_2+16\gamma_1^2+4\gamma_4-9\gamma_2^2)\right]_{-2}^{1}$$
$$+\left[\frac{1}{24r^3}(-4-84\gamma_2-160\gamma_1^2-40\gamma_4+48\gamma_2^2-96\gamma_3\gamma_1-3\gamma_6+96\gamma_2\gamma_1^2+24\gamma_4\gamma_2-30\gamma_2^3)\right]_{-2}^{1},$$

$$\kappa_2(z) = \frac{1}{4}\left[\frac{1}{r}(2+\gamma_2)\right]_{+2}^{1} + \frac{1}{8}\left[\frac{1}{r^2}(-8\gamma_2-2\gamma_4-8\gamma_1^2+5\gamma_2^2)\right]_{+2}^{1}$$
$$+\frac{1}{48}\left[\frac{1}{r^3}(-8+168\gamma_2+120\gamma_4+480\gamma_1^2-152\gamma_2^2+352\gamma_3\gamma_1\right.$$
$$\left.+11\gamma_6-96\gamma_4\gamma_2-384\gamma_2\gamma_1^2+128\gamma_2^3)\right]_{+2}^{1},$$

$$\kappa_3(z) = \frac{1}{8}\left[\frac{1}{r^2}(-4+4\gamma_1^2+\gamma_4-3\gamma_2^2)\right]_{-2}^{1}$$
$$+\frac{1}{16}\left[\frac{1}{r^3}(-80\gamma_1^2-20\gamma_4+24\gamma_2^2-96\gamma_3\gamma_1-3\gamma_6+120\gamma_2\gamma_1^2+30\gamma_4\gamma_2-44\gamma_2^3)\right]_{-2}^{1},$$

$$\kappa_4(z) = \frac{1}{16}\left[\frac{1}{r^3}(16+8\gamma_2^2+32\gamma_3\gamma_1-48\gamma_2\gamma_1^2+\gamma_6-12\gamma_4\gamma_2+20\gamma_2^3)\right]_{+2}^{1}.$$

For cumulants of z for samples from a normal population we have, to order $1/r^3$,

$$\mathscr{E}(z) = -\frac{1}{2r_1}+\frac{1}{2r_2}+\frac{1}{3r_1^2}-\frac{1}{3r_2^2}-\frac{1}{6r_1^3}+\frac{1}{6r_2^3},$$

$$\kappa_2(z) = \frac{1}{2r_1}+\frac{1}{2r_2}-\frac{1}{6r_1^3}-\frac{1}{6r_2^3},$$

$$\kappa_3(z) = -\frac{1}{2r_1^2}+\frac{1}{2r_2^2},$$

$$\kappa_4(z) = \frac{1}{r_1^3}+\frac{1}{r_2^3}.$$

Table 2 gives a comparison of the true values, from Wishart, with those obtained by substituting in the formulae for normal parents immediately above. The case considered is $f_1 = 24, f_2 = 60$. The agreement is satisfactory.

Table 2. *Cumulant constants of the distribution of z ($f_1 = 24, f_2 = 60$; parent population normal)*

	κ_1	κ_2	κ_3	κ_4
Wishart's exact values	−0·0127,429	0·0301,992	−0·0007,998	0·0000,867
Approximate values	−0·0127,431	0·0301,992	−0·0008,015	0·0000,871

8. Geary, in his 1947 paper, has discussed the effect of kurtosis on the distribution of z when both samples are drawn from the same population. As an illustration of the cumulants of z derived here we shall discuss the effect of skewness in the parent population for the case $f_1 = 24, f_2 = 60$. It will be assumed that the parent population may be graduated by the first three terms of the Gram-Charlier Type A series,* i.e. that

$$y = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}X^2}\left(1 + \frac{\gamma_1}{3!}H_3(X) + \frac{\gamma_2}{4!}H_4(X)\right),$$

where $H_3(X)$ and $H_4(X)$ are the third and fourth Hermite polynomials. This parent population has cumulants

$$\kappa_1,\ \kappa_2,\ \kappa_3,\ \kappa_4,\ \kappa_5 = 0, \quad \kappa_6 = -10\kappa_3^2, \quad \kappa_7 = 35\kappa_4\kappa_3, \quad \kappa_8 = -35\kappa_4^2.$$

In order to eliminate as far as possible any effect of kurtosis we shall put $\kappa_4 = 0$, when the only higher cumulant which has a non-zero value will be $\kappa_6 = -10\kappa_3^2$. Under this assumption and assuming further that $\kappa_2 = 1$ the momental constants of the distribution of z for different degrees of skewness are as given in Table 3. $\kappa_4(z|\kappa_3)$ is unaltered at the degree of approximation to which we are working. It would appear from a study of these moments that the effect of skewness on the distribution of z is likely to be small, and this is, in fact, the case. There are

Table 3. *Momental constants of z ($f_1 = 24, f_2 = 60$) when κ_4 of parent population is zero*

κ_3^2	0·0	0·1	0·3	0·5
$\mu_1'(z\|\kappa_3)$	−0·012,743	−0·012,826	−0·012,992	−0·013,158
$\mu_2(z\|\kappa_3)$	0·030,199	0·030,395	0·030,787	0·031,179
$\sigma(z\|\kappa_3)$	0·173,779	0·174,342	0·175,462	0·176,576
$\kappa_3(z\|\kappa_3)$	−0·000,801	−0·000,864	−0·000,988	−0·001,113

various ways in which we can estimate the effect of this skewness. We could use the moments of $z|\kappa_3$ to find the Pearson curve with the appropriate β's, or follow the Fisher-Cornish procedure and use Edgeworth's series or fit a Gram-Charlier Type A of the functional form given above. This last procedure is approximate but will be adequate for our purposes. Thus if

* I owe to Dr J. Wishart the suggestion that a more appropriate form for the population would be

$$y = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}X^2}\left(1 + \gamma_1\frac{H_3(X)}{3!} + \gamma_2\frac{H_4(X)}{4!} + \frac{10\gamma_1^2 H_6(X)}{6!}\right).$$

I agree that for any systematic investigation of the effect of skewness in the parent population on the distribution of z it might be better to take this functional form. In the case above, however, I am specifying a particular population and my purpose is one of illustration only.

$z_{0\cdot05}$ is the deviate beyond which 5 % of the frequency might be expected to lie when κ_3 of the parent population is zero, it is seen that we require to evaluate

$$\Phi = \frac{1}{\sqrt{(2\pi)}}\int_{\frac{z_{0\cdot05}-\kappa_1(z|\kappa_3)}{\sqrt{[\kappa_2(z|\kappa_3)]}}}^{\infty} e^{-\frac{1}{2}X^2}dX + \frac{\gamma_1(z|\kappa_3)}{3!}\int_{\frac{z_{0\cdot05}-\kappa_1(z|\kappa_3)}{\sqrt{[\kappa_2(z|\kappa_3)]}}}^{\infty} H_3(X)\frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}X^2}dX$$
$$+\frac{\gamma_2(z|\kappa_3)}{4!}\int_{\frac{z_{0\cdot05}-\kappa_1(z|\kappa_3)}{\sqrt{[\kappa_2(z|\kappa_3)]}}}^{\infty} H_4(X)\frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}X^2}dX,$$

remembering that

$$\frac{1}{\sqrt{(2\pi)}}\int_X^{+\infty} H_r(X)\,e^{-\frac{1}{2}X^2}dX = \frac{1}{\sqrt{(2\pi)}}\Big[H_{r-1}(X)\,e^{-\frac{1}{2}X^2}\Big]_\infty^X.$$

9. As a check on the adequacy of the Gram-Charlier Type A we first refer to the tabled values of z, finding that

$$z_{0\cdot05} = 0\cdot26535 \quad \text{for} \quad \kappa_3 = 0,\ f_1 = 24,\ f_2 = 60,$$

and then proceed to find the tail area cut off by $z_{0\cdot05}$ for these values using the expansion just above. We find

$$\Phi(\kappa_3 = 0) = 0\cdot05477 - 0\cdot00439 - 0\cdot00031 = 0\cdot05007.$$

Thus the representation of the z distribution by the normal curve plus the first two corrective terms of the Gram-Charlier A gives the probability integral correct to three decimal places, which will mean that it gives sufficient accuracy for our purpose. We repeat the procedure for $\kappa_3^2 = 0\cdot1, 0\cdot3, 0\cdot5$ and draw up Table 4. It is clear that a moderate amount of skewness in the parent population will not affect the z test appreciably for the case considered ($f_1 = 24$, $f_2 = 60$), and this will be true for higher values of f_1 and f_2.

Table 4. *Tail areas corresponding to* $z_{0\cdot05}$ $(\kappa_3 = 0)$ *when* $\kappa_3 \neq 0$.

κ_3^2	0·0	0·1	0·3	0·5
Φ	0·050	0·050	0·051	0·051

10. To investigate the effect of skewness in the parent population for degrees of freedom less than those taken in the preceding paragraph, the procedure is complicated a little by the fact that the cumulants of z are of order $1/n^3$ only, and that the Gram-Charlier A does not give exact results. It is perhaps useful, however, to give such calculations as were carried out. We consider the case $f_1 = 8$, $f_2 = 16$, and take only the normal curve and the third Hermite polynomial to represent the distribution of z. Carrying through the calculations as before we may draw up Table 5. For $\kappa_3^2 = 0$, Φ should be equal to 0·050 and not 0·051, the

Table 5. *Tail areas cut off by* $z_{0\cdot05}$ $(= 0\cdot4760)$, $f_1 = 8, f_2 = 16$

κ_3^2	0·0	0·1	0·3	0·5
Φ	0·051	0·052	0·053	0·054

inaccuracy being introduced partly by the moments and partly by the expansion. However, the figures are, I think, sufficient to show that for the case considered a small amount of skewness in the parent population will not affect the z test very much.

11. The moments of F can be obtained by precisely the same method as was used for obtaining the moments of z, although the results are not as satisfactory and the algebraic manipulation required is very much heavier. We have

$$F = \frac{k_2'}{k_2} = \frac{\kappa_2'\left(1+\frac{k_2'-\kappa_2'}{\kappa_2'}\right)}{\kappa_2\left(1+\frac{k_2-\kappa_2}{\kappa_2}\right)},$$

the dashes being given to the numerator this time to avoid much repetition. Thus

$$\frac{\kappa_2}{\kappa_2'}\mathscr{E}(F) = \left[1+\frac{1}{\kappa_2^2}(\overline{2^2})-\frac{1}{\kappa_2^3}(\overline{2^3})+\frac{1}{\kappa_2^4}(\overline{2^4})-\frac{1}{\kappa_2^5}(\overline{2^5})+\dots\right],$$

and on substitution we have, reducing as before,

$$\begin{aligned}\mathscr{E}(F) = \frac{\kappa_2'}{\kappa_2}\Bigg[1&+\frac{2}{f_2}+\frac{4}{f_2^2}+\frac{8}{f_2^3}+\gamma_2\left(\frac{1}{f_2}-\frac{1}{f_2^2}+\frac{5}{f_2^3}\right)-\gamma_1^2\left(\frac{1}{f_2^2}+\frac{2}{f_2^3}\right)\\&-\gamma_3\left(\frac{1}{f_2^2}-\frac{6}{f_2^3}\right)+\gamma_2^2\left(\frac{3}{f_2^2}-\frac{4}{f_2^3}\right)+\gamma_3\gamma_1\frac{32}{f_2^3}+\gamma_6\frac{1}{f_2^3}-\gamma_2\gamma_1^2\frac{40}{f_2^3}-\gamma_4\gamma_2\frac{10}{f_2^3}+\gamma_2^3\frac{15}{f_2^3}\Bigg].\end{aligned}$$

Collecting up terms we may write

$$\begin{aligned}\mathscr{E}(F) = \frac{\kappa_2'}{\kappa_2}\Bigg[1&+\frac{1}{f_2}(2+\gamma_2)+\frac{1}{f_2^2}(4-\gamma_2-\gamma_1^2-\gamma_3+3\gamma_2^2)\\&+\frac{1}{f_2^3}(8+5\gamma_2-2\gamma_1^2+6\gamma_3-4\gamma_2^2+32\gamma_3\gamma_1+\gamma_6-40\gamma_2\gamma_1^2-10\gamma_4\gamma_2+15\gamma_2^3)\Bigg].\end{aligned}$$

Similarly by expanding $\mathscr{E}(F^2)$ and $\mathscr{E}(F^3)$, correcting to moments about the mean, and collecting terms we obtain

$$\begin{aligned}\sigma_F^2 = \frac{\kappa_2'^2}{\kappa_2^2}\Bigg[&\frac{1}{f_1}(2+\gamma_2')+\frac{1}{f_2}(2+\gamma_2)-\frac{\gamma_2'}{f_1^2}+\frac{1}{f_2^2}(16+7\gamma_2-8\gamma_1^2-2\gamma_4+8\gamma_2^2)+\frac{3}{f_1f_2}(2+\gamma_2')(2+\gamma_2)\\&+\frac{\gamma_2'}{f_1^2}-\frac{3\gamma_2'(2+\gamma_2)}{f_1^2f_2}+\frac{(2+\gamma_2')}{f_1f_2^2}(28+9\gamma_2-16\gamma_1^2-4\gamma_3+15\gamma_2^2)\\&+\frac{1}{f_2^3}(88+37\gamma_2-32\gamma_1^2+38\gamma_2^2+96\gamma_3\gamma_1+3\gamma_6-152\gamma_2\gamma_1^2-38\gamma_4\gamma_2+69\gamma_2^3)\Bigg],\end{aligned}$$

$$\begin{aligned}\mu_3(F) = \frac{\kappa_2'^3}{\kappa_2^3}\Bigg[&\frac{1}{f_1^2}(\gamma_4'+4\gamma_1'^2+12\gamma_2'+8)+\frac{6}{f_1f_2}(2+\gamma_2)(2+\gamma_2')+\frac{1}{f_2^2}(16+12\gamma_2-4\gamma_1^2-\gamma_4+6\gamma_2^2)\\&+\frac{1}{f_1^3}(-2\gamma_4'-8\gamma_1'^2-12\gamma_2')+\frac{6(2+\gamma_2)}{f_1^2f_2}(\gamma_4'+4\gamma_1'^2+11\gamma_2'+8)\\&+\frac{3}{f_1f_2^2}(2+\gamma_2')(56+34\gamma_2-20\gamma_1^2-5\gamma_4+24\gamma_2^2)\\&+\frac{1}{f_2^3}(256+180\gamma_2-208\gamma_1^2-28\gamma_4+168\gamma_2^2+96\gamma_3\gamma_1+3\gamma_6-204\gamma_2\gamma_1^2-51\gamma_4\gamma_2+116\gamma_2^3)\Bigg].\end{aligned}$$

$\mu_4(F)$ is of great length and complexity, and for this and reasons given in the succeeding paragraphs we have not written it out here.

12. It is clear that except for large values of n_1 and n_2, the expansion gives expressions which are not very good approximations to the true moments. Moreover, it is not thought that retaining higher powers of $1/f$ in the expansion will help very much. The numerical coefficients are increasing so rapidly that for small f the quasi-asymptotic series begin to diverge before they have converged to any quantity close to the true value. This is perhaps most easily seen if we consider the case of the moments of F when the two samples have been drawn from the same normal population. For this case the true moments are easily found to be (see, for example, Wishart (1947*b*)),

$$\mu_1(F) = \frac{f_2}{f_2-2}, \quad \mu_2(F) = \frac{2f_2^2(f_1+f_2-2)}{f_1(f_2-2)^2(f_2-4)}, \quad \mu_3(F) = \left(\frac{f_2}{f_1}\right)^3 \frac{8f_1(f_1+f_2-2)(2f_1+f_2-2)}{(f_2-2)^3(f_2-4)(f_2-6)}.$$

In the expansion in § 11 we write $\kappa_2 = \kappa_2'$ and put all the other κ's equal to zero. We have then

$$\mathscr{E}(F) = 1+\frac{2}{f_2}+\frac{4}{f_2^2}+\frac{8}{f_2^3},$$

$$\sigma_F^2 = 2\left(\frac{1}{f_1}+\frac{1}{f_2}\right)+\frac{12}{f_1f_2}+\frac{16}{f_2^2}+\frac{56}{f_1f_2^2}+\frac{88}{f_2^3},$$

$$\mu_3(F) = \frac{8}{f_1^2}+\frac{24}{f_1f_2}+\frac{16}{f_2^2}+\frac{96}{f_1^2f_2}+\frac{336}{f_1f_2^2}+\frac{256}{f_2^3}.$$

These approximate moments may be checked by expanding the true moments as power series in $1/f_1$, and $1/f_2$, whence the leading terms are found to agree with the approximate values just given. Numerical substitution shows, however, that these approximate expressions do not agree very well with the true values even for $f_1 = f_2 = 20$.

13. Possibly a better approximation to the moments of F when the samples are from any two parent populations may be obtained by the following artifice. The moments of F when the samples are both from the same normal population are given by those terms inside the brackets which are not multiplied by any γ's. We know (Wishart, 1947*b*) the exact values $\mu_1'(F_N)$, $\mu_2(F_N)$, $\mu_3(F_N)$, for which these values are only an approximation. Let us write, therefore,

$$\mathscr{E}(F) = \frac{\kappa_2'}{\kappa_2}\left[\mu_1'(F_N)+\frac{\gamma_2}{f_2}-\frac{1}{f_2^2}(\gamma_2+\gamma_1^2+\gamma_3-3\gamma_2^2)\right.$$
$$\left.+\frac{1}{f_2^3}(5\gamma_2-2\gamma_1^2+6\gamma_3-4\gamma_2^2+32\gamma_3\gamma_1+\gamma_6-40\gamma_2\gamma_1^2-10\gamma_4\gamma_2+15\gamma_2^3)\right],$$

$$\sigma_F^2 = \frac{\kappa_2'^2}{\kappa_2^2}\left[\mu_2(F_N)+\frac{\gamma_2'}{f_1}+\frac{\gamma_2}{f_2}-\frac{\gamma_2'}{f_1^2}+\frac{1}{f_2^2}(7\gamma_2-8\gamma_1^2-2\gamma_4+8\gamma_2^2)+\frac{3}{f_1f_2}(\gamma_2'(2+\gamma_2)+\gamma_2(2+\gamma_2'))\right.$$
$$+\frac{\gamma_2'}{f_1^2}-\frac{3\gamma_2'(2+\gamma_2)}{f_1^2f_2}+\frac{(2+\gamma_2')(28+9\gamma_2-16\gamma_1^2-4\gamma_3+15\gamma_2^2)-56}{f_1f_2^2}$$
$$\left.+\frac{1}{f_2^3}(37\gamma_2-32\gamma_1^2+38\gamma_2^2+96\gamma_3\gamma_1+3\gamma_6-152\gamma_2\gamma_1^2-38\gamma_4\gamma_2+69\gamma_2^3)\right],$$

$$\mu_3(F) = \frac{\kappa_2'^3}{\kappa_2^3}\left[\mu_3(F_N)+\frac{1}{f_1^2}(\gamma_4'+4\gamma_1'^2+12\gamma_2')+\frac{6(2+\gamma_2)(2+\gamma_2')-24}{f_1f_2}+\frac{1}{f_2^2}(12\gamma_2-4\gamma_1^2-\gamma_4+6\gamma_2^2)\right.$$
$$-\frac{1}{f_1^3}(2\gamma_4'+8\gamma_1'^2+12\gamma_2')+\frac{6(2+\gamma_2)(\gamma_4'+4\gamma_1'^2+11\gamma_2'+8)-96}{f_1^2f_2}$$
$$+\frac{3(2+\gamma_2')(56+34\gamma_2-20\gamma_1^2-5\gamma_4+24\gamma_2^2)-336}{f_1f_2^2}$$
$$\left.+\frac{1}{f_2^3}(180\gamma_2-208\gamma_1^2-28\gamma_4+168\gamma_2^2+96\gamma_3\gamma_1+3\gamma_6-204\gamma_2\gamma_1^2-51\gamma_4\gamma_2+116\gamma_2^3)\right].$$

In these expansions the γ's play the role of terms which are correcting the true normal moments for the departure from normality. Provided γ_1 and γ_2 are less than unity these expressions should give reasonable approximations to the true moments of F when the samples are drawn from any two parent populations whatever. It will be noted, however, that in $\mu_3(F)$ the numerical coefficients are rather large, indicating that the approximation will not be too good. This remark holds good *a fortiori* for $\mu_4(F)$.

I should like to thank Prof. E. S. Pearson and Dr J. Wishart for helpful criticism.

REFERENCES

CORNISH, E. A. & FISHER, R. A. (1937). *Revue de l'Institut International de Statistique*, **5**, 307.
DAVID, F. N. (1949). *Biometrika*, **36**, 383.
DERKSEN, J. B. D. (1939). *Ann. Math. Statist.* **10**, 380.
GEARY, R. C. (1947). *Biometrika*, **34**, 209.
PEARSON, E. S. (1931). *Biometrika*, **23**, 119.
WISHART, J. (1947*a*). *Biometrika*, **34**, 170.
WISHART, J. (1947*b*). *J. Inst. Actuar. Stud. Soc.* **6**, 172.

THE METHOD OF FREQUENCY-MOMENTS AND ITS APPLICATION TO TYPE VII POPULATIONS

By HERBERT S. SICHEL

National Institute for Personnel Research, South African Council for Scientific and Industrial Research

Part 1. Theoretical

The method of frequency-moments was developed primarily with the object of overcoming certain difficulties when fitting growth curves to observed data. In the author's original paper (1947) examples were given on how to fit exponential, logistic and Gompertz functions. In the second part of the same paper it was pointed out that the method could also be applied to the graduation of frequency-distributions and examples of normal, lognormal and Pearson Type VII curves were presented.

Yule's (1938) investigation, based on an early suggestion of Karl Pearson, has come to the author's notice recently. The fitting process described by Yule is essentially the same as my method of frequency-moments. The derivation of the standard errors given on the following pages is based partly on Yule's work.

In this paper we are more interested in the problem of estimating the parameters of a given type of frequency-distribution by the method of frequency-moments than in fitting the proposed distribution to a set of observations, though this aspect is dealt with in the last section. The present investigation has been limited to the case of frequency-distributions only, and the section comparing the maximum likelihood solution with the frequency-moment method has been confined to a Pearson Type VII population.

Definition of parameters

The nth frequency-moment of a population represented by a probability law

$$y\,dx = f(x)\,dx$$

will be defined as

$$J_n = N^n \int_{-\infty}^{+\infty} y^n dx, \tag{1}$$

where N is the total number of items in the sample and is equal to J_1. Corresponding to (1) we shall define the nth probability-moment of the population as

$$\Omega_n = \frac{J_n}{J_1^n} = \int_{-\infty}^{+\infty} y^n dx. \tag{2}$$

In practice we are rarely in the position of estimating Ω_n directly. However, it is comparatively easy to estimate the parameter

$$\omega_n = \left(\int_{-\infty}^{a} y\,dx\right)^n + \sum_{i=1}^{r}\left(\int_{a+\frac{b-a}{r}(i-1)}^{a+\frac{b-a}{r}i} y\,dx\right)^n + \left(\int_{b}^{+\infty} y\,dx\right)^n = \sum_{i=0}^{r+1} \pi_i^n, \text{ say.} \tag{3}$$

By suitable choice of a and b we can make

$$\pi_0 = \int_{-\infty}^{a} y\,dx \quad \text{and} \quad \pi_{r+1} = \int_{b}^{+\infty} y\,dx$$

smaller than any preassigned ϵ, so that in practice

$$\omega_n \doteqdot \sum_{i=1}^{r} \left(\int_{a+\frac{b-a}{r}(i-1)}^{a+\frac{b-a}{r}i} y\,dx \right)^n = \sum_{i=1}^{r} \pi_i^n, \tag{3a}$$

where r is the number of classes of equal width $h = (b-a)/r$ into which the population has been subdivided.

In the limit we have $\lim_{h \to 0} h^{1-n}\,\omega_n = \Omega_n$.

For computational work it is convenient to make h equal to unity. Even for r as small as 15 the approximation

$$h^{1-n}\omega_n \doteqdot \Omega_n$$

is reasonable for most cases provided n is of a low order. It may be improved by the use of a correction to be described in Part 2. The parameter ω_n is hereafter called the nth working probability-moment.

Certain ratios of working probability-moments are defined as

$$\left.\begin{aligned} \alpha_n &= \frac{\omega_{n+2}}{\omega_2^{n+1}} \quad (n = 1, 2, 3, 4, \ldots), \\ \text{and} \quad \alpha_n &= \frac{\omega_{\frac{1}{2}(2n+5)}}{\omega_1^{\frac{1}{2}(2n+1)}} \quad (n = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \tfrac{7}{2}, \ldots). \end{aligned}\right\} \tag{4}$$

The α_n coefficients may be used as measures of kurtosis.

The dispersion of a population may be represented by parameters of the nature

$$\rho_{n-1} = \frac{1}{\omega_n}, \tag{5}$$

subject to certain limitations mentioned in Yule's (1938) paper.

When scale and location of a distribution are changed we transform the variate by

$$z = kx + l.$$

The transformed distribution becomes

$$p(z) = \frac{1}{k} f\left(\frac{z-l}{k}\right)$$

and

$$\omega_n(z) = \sum_{i=0}^{r+1} \left\{ \frac{1}{k} \int_{z_i}^{z_{i+1}} f\left(\frac{z-l}{k}\right) dz \right\}^n,$$

$$\omega_n(z) = k^{1-n}\omega_n. \tag{6}$$

It follows that the working probability-moments are independent of the constant l and have to be multiplied by k^{1-n} if the variate is multiplied by k. They are, therefore, semi-invariant under the transformation $z = kx + l$.

The α_n coefficients are unaffected by the transformation, as can easily be shown by substitution of equation (6) into (4).

In the following, extensive use will be made of working *probability*-moments. In practical applications, however, we usually deal with working *frequency*-moments. In this fact lies the justification of the name suggested for the new method.

Large sample mean values, variances and covariances of statistics o_n and a_n

We may estimate ω_n by
$$o_n = \sum_{i=0}^{r+1} p_i^n, \tag{7}$$
where $p_i = f_i/N$ = observed proportion of observations falling into the ith class interval. Denoting the deviation of p_i from its mean value π_i by
$$\delta p_i = p_i - \pi_i,$$
we have for the mean value of o_n
$$\begin{aligned} E(o_n) &= \Sigma E(p_i^n) = \Sigma E(\pi_i + \delta p_i)^n \\ &= \Sigma E(\pi_i^n + n\pi_i^{n-1}\delta p_i + n_{(2)}\pi_i^{n-2}\delta p_i^2 + \ldots), \end{aligned}$$
where $n_{(r)}$ is written for $\binom{n}{r}$. Neglecting higher powers of δp_i as $E(\delta p_i^3)$ will be of order N^{-2}, and using results obtained from the binomial theorem
$$E(\delta p_i) = 0 \quad \text{and} \quad E(\delta p_i^2) = \frac{\pi_i(1-\pi_i)}{N},$$
we find
$$\begin{aligned} E(o_n) &= \Sigma\left[\pi_i^n + \frac{n_{(2)}}{N}(\pi_i^{n-1} - \pi_i^n)\right] + O(N^{-2}) \\ &= \omega_n + \frac{n_{(2)}}{N}(\omega_{n-1} - \omega_n) + O(N^{-2}). \end{aligned} \tag{8}$$
For large N
$$E(o_n) = \omega_n + O(N^{-1}). \tag{9}$$
Writing δo_n for the deviation of o_n from its mean value $E(o_n)$, we have
$$\begin{aligned} \delta o_n &= o_n - E(o_n) = \Sigma p_i^n - \Sigma \pi_i^n + O(N^{-1}) \\ &= \Sigma[n\pi_i^{n-1}\delta p_i + n_{(2)}\pi_i^{n-2}\delta p_i^2] + O(N^{-1}). \end{aligned}$$
After squaring this expression and taking mean values we have
$$\operatorname{var}(o_n) = n^2[\Sigma\pi_i^{2n-2}E(\delta p_i^2) + \Sigma(\pi_i\pi_j)^{n-1}E(\delta p_i \delta p_j)] + O(N^{-2}),$$
where $i \neq j$ and all permutations are permitted. On substituting the value for $E(\delta p_i^2)$ given previously and also the covariance
$$E(\delta p_i \delta p_j) = -\frac{\pi_i \pi_j}{N},$$
we obtain
$$\begin{aligned} \operatorname{var}(o_n) &= \frac{n^2}{N}[\Sigma\pi_i^{2n-1} - \Sigma\pi_i^{2n} - \Sigma(\pi_i\pi_j)^n] + O(N^{-2}) \\ &= \frac{n^2}{N}[\omega_{2n-1} - \omega_n^2] + O(N^{-2}). \end{aligned} \tag{10}$$
As an example, let us find the efficiency of the method of frequency-moments in estimating the standard deviation σ of a normal population. For class width $h = 1$
$$\begin{aligned} \omega_n \doteq \Omega_n &= \left(\frac{1}{\sigma\sqrt{(2\pi)}}\right)^n \int_{-\infty}^{+\infty} e^{-nx^2/2\sigma^2}\,dx \\ &= (2\pi\sigma^2)^{\frac{1}{2}(1-n)}n^{-\frac{1}{2}}. \end{aligned}$$

σ, expressed in terms of working probability-moments of order $\frac{3}{2}$ ($n = \frac{3}{2}$ for reasons to be explained at a later stage; see equation (53)), is

$$\sigma = \frac{1}{3}\sqrt{\frac{2}{\pi}}\omega_{\frac{1}{2}}^{-2}.$$

We can, therefore, estimate σ by $$s' = \frac{1}{3}\sqrt{\frac{2}{\pi}}o_{\frac{1}{2}}^{-2}. \tag{11}$$

Writing (11) as $$s' = \frac{1}{3}\sqrt{\frac{2}{\pi}}[E(o_{\frac{1}{2}}) + \delta o_{\frac{1}{2}}]^{-2},$$

and expanding the bracket for small deviations of $\delta o_{\frac{1}{2}}$ such that

$$\left|\frac{\delta o_{\frac{1}{2}}}{E(o_{\frac{1}{2}})}\right| < 1,$$

$$s' = \frac{1}{3}\sqrt{\frac{2}{\pi}}[E(o_{\frac{1}{2}})]^{-2}\left[1 - 2\frac{\delta o_{\frac{1}{2}}}{E(o_{\frac{1}{2}})} + - \ldots\right], \quad E(s') = \frac{1}{3}\sqrt{\frac{2}{\pi}}\omega_{\frac{1}{2}}^{-2} + O(N^{-1}). \tag{12}$$

It follows from (12) that s' is a consistent estimator.

Further, for small $\delta o_{\frac{1}{2}}$

$$\text{var}\,(s') = \frac{8}{9\pi}\omega_{\frac{1}{2}}^{-6}\,\text{var}\,(o_{\frac{1}{2}}),$$

and hence, from (10), $$\text{var}\,(s') = \frac{2}{N\pi}\omega_{\frac{1}{2}}^{-6}[\omega_2 - \omega_{\frac{1}{2}}^2]$$

$$= \frac{9}{N}(\tfrac{3}{4}\sqrt{2} - 1)\sigma^2. \tag{13}$$

Hence the efficiency of s' is given for large N by

$$\text{Eff.}\,(s') = \frac{1}{9\left(\frac{3}{\sqrt{2}} - 2\right)} = 0{\cdot}916, \tag{14}$$

which is good in comparison with the efficiencies of other estimators sometimes used in place of the moment-statistics.

The covariances of working probability-moments are derived in a similar way to the variances. We find

$$\text{cov}\,(o_n, o_m) = \frac{nm}{N}(\omega_{n+m-1} - \omega_n\omega_m) + O(N^{-2}). \tag{15}$$

The α_n ratios of the working probability-moments may be estimated by

$$\left.\begin{aligned} a_n &= \frac{o_{n+2}}{o_2^{n+1}} \quad (n = 1, 2, 3, 4, \ldots), \\ \text{and} \qquad a_n &= \frac{o_{\frac{1}{2}(2n+5)}}{o_{\frac{1}{2}}^{\frac{1}{2}(2n+1)}} \quad (n = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \tfrac{7}{2}, \ldots). \end{aligned}\right\} \tag{16}$$

Now if $a_n = \phi(o_k, o_l)$, we find in the usual way

$$\text{var}\,(a_n) = \left(\frac{\partial\phi}{\partial o_k}\right)^2 \text{var}\,(o_k) + 2\frac{\partial\phi}{\partial o_k}\frac{\partial\phi}{\partial o_l}\,\text{cov}\,(o_k, o_l) + \left(\frac{\partial\phi}{\partial o_l}\right)^2 \text{var}\,(o_l). \tag{17}$$

This result is correct to order N^{-1}. For $n = \frac{3}{2}$ we find

$$\text{var}\,(a_1) = \frac{1}{N}[4\alpha_1 + 9\alpha_1^3 - \alpha_1^2 - 12\alpha_1\alpha_1] + O(N^{-2}), \tag{18}$$

and for $n = 1$

$$\text{var}\,(a_1) = \frac{1}{N}[9\alpha_3 + 16\alpha_1^3 - \alpha_1^2 - 24\alpha_1\alpha_2] + O(N^{-2}). \tag{19}$$

First three exact moments of statistic o_2

With a view of getting some idea as to the sample size N for which one may expect the statistics o_2 and o_1 to be normally distributed, it was originally intended to derive the first four exact moments of o_2. It was assumed that if for a given N the statistic o_2 is nearly normal it would not be unreasonable to expect o_1, being of lower order, also to be normally distributed for the same N. The algebra, however, was found to be extremely heavy. In the following the first three exact moments of o_2 are given only. The fourth moment of o_2, derived below, is correct to order N^{-2}.

In general we have

$$E(o_n) = \Sigma E(\pi_i + \delta p_i)^n = \Sigma E(\pi_i^n + n\pi_i^{n-1}\delta p_i + n_{(2)}\pi_i^{n-2}\delta p_i^2 + \ldots), \tag{20}$$

and

$$\begin{aligned}\delta o_n &= o_n - E(o_n)\\ &= n\,\Sigma\pi_i^{n-1}[\delta p_i - E(\delta p_i)] + n_{(2)}\,\Sigma\pi_i^{n-2}[\delta p_i^2 - E(\delta p_i^2)] + n_{(3)}\,\Sigma\pi_i^{n-3}[\delta p_i^3 - E(\delta p_i^3)] + \ldots.\end{aligned} \tag{21}$$

In particular,

$$\mu_1'(o_2) = E(o_2) = \omega_2 + \frac{1}{N}(1 - \omega_2), \tag{22}$$

being the exact mean value of o_2, and

$$\delta o_2 = 2\,\Sigma\pi_i\delta p_i + \Sigma\delta p_i^2 + \frac{\omega_2 - 1}{N}.$$

Hence the exact nth moment of o_2

$$\mu_n(o_2) = E(\delta o_2^n) = E\left(2\,\Sigma\pi_i\delta p_i + \Sigma\delta p_i^2 + \frac{\omega_2 - 1}{N}\right)^n. \tag{23}$$

For the second moment we find

$$\begin{aligned}\mu_2(o_2) = {} & \Sigma E(\delta p_i^4) + \Sigma E(\delta p_i^2\delta p_j^2) + 4\,\Sigma\pi_i E(\delta p_i^3) + 4\,\Sigma\pi_i E(\delta p_i\delta p_j^3) + 4\,\Sigma\pi_i^2 E(\delta p_i^3)\\ & + 4\,\Sigma\pi_i\pi_j E(\delta p_i\delta p_j) + \frac{2(\omega_2 - 1)}{N}\Sigma E(\delta p_i^2) + \frac{(\omega_2 - 1)^2}{N^2},\end{aligned} \tag{24}$$

where $i \neq j$ and all permutations permitted. In order to evaluate (24) we must know the central bivariate moments

$$\mu_{k_1k_2} = E(\delta p_i^{k_1}\delta p_j^{k_2}) = \frac{1}{N^{k_1+k_2}}E(\delta f_i^{k_1}\delta f_j^{k_2})$$

of the multinomial distribution

$$P(f_1, f_2, f_3, \ldots, f_r) = \frac{N!}{f_1!\,f_2!\,f_3!\ldots f_r!}\pi_1^{f_1}\pi_2^{f_2}\pi_3^{f_3}\ldots\pi_r^{f_r}.$$

In general we require for the derivation of $\mu_n(o_2)$ central multivariate moments (n variables)

$$\begin{aligned}\mu_{k_1k_2k_3\ldots k_n} &= E(\delta p_{i_1}^{k_1}\delta p_{i_2}^{k_2}\delta p_{i_3}^{k_3}\ldots\delta p_{i_n}^{k_n})\\ &= \frac{1}{N^{k_1+k_2+k_3+\ldots+k_n}}E(\delta f_{i_1}^{k_1}\delta f_{i_2}^{k_2}\delta f_{i_3}^{k_3}\ldots\delta f_{i_n}^{k_n}),\end{aligned} \tag{25}$$

where the various l's can take on all the values 1, 2, 3, ..., r, r being the number of classes into which the population has been subdivided. The order of the multivariate moment is

$$k_1+k_2+k_3+\ldots+k_n.$$

The joint moment-generating function of the multinomial distribution for the case of n variables is

$$M\begin{pmatrix} f_{l_1} f_{l_2} f_{l_3} \cdots f_{l_n} \\ t_{l_1} t_{l_2} t_{l_3} \cdots t_{l_n} \end{pmatrix} = (\pi_{l_1} e^{t_{l_1}} + \pi_{l_2} e^{t_{l_2}} + \pi_{l_3} e^{t_{l_3}} + \ldots + \pi_{l_n} e^{t_{l_n}} + \pi)^N, \tag{26}$$

where $$\pi = 1-\pi_{l_1}-\pi_{l_2}-\pi_{l_3}-\ldots-\pi_{l_n}.$$

The expansion of this expression and the collection of appropriate terms of $t_{l_n}^{k_n}/k_n!$ is straightforward but laborious. The various multivariate moments derived from (26) are moments about the origin with respect to variables $f_{l_n} = Np_{l_n}$. They are denoted as $\nu'_{k_1 k_2 k_3 \ldots k_n}$. The following moments are needed for subsequent work. Writing i, j, k, l for l_1, l_2, l_3, l_4 and $N^{(r)} = N(N-1)\ldots(N-r+1)$, we find

$$\left.\begin{aligned}
\nu'_{1000} &= N\pi_i, \\
\nu'_{2000} &= N\pi_i + N^{(2)}\pi_i^2, \\
\nu'_{3000} &= N\pi_i + 3N^{(2)}\pi_i^2 + N^{(3)}\pi_i^3, \\
\nu'_{4000} &= N\pi_i + 7N^{(2)}\pi_i^2 + 6N^{(3)}\pi_i^3 + N^{(4)}\pi_i^4, \\
\nu'_{5000} &= N\pi_i + 15N^{(2)}\pi_i^2 + 25N^{(3)}\pi_i^3 + 10N^{(4)}\pi_i^4 + N^{(5)}\pi_i^5, \\
\nu'_{6000} &= N\pi_i + 31N^{(2)}\pi_i^2 + 90N^{(3)}\pi_i^3 + 65N^{(4)}\pi_i^4 + 15N^{(5)}\pi_i^5 + N^{(6)}\pi_i^6, \\
\nu'_{1100} &= N^{(2)}\pi_i\pi_j, \\
\nu'_{1200} &= N^{(2)}\pi_i\pi_j + N^{(3)}\pi_i\pi_j^2, \\
\nu'_{1300} &= N^{(2)}\pi_i\pi_j + 3N^{(3)}\pi_i\pi_j^2 + N^{(4)}\pi_i\pi_j^3, \\
\nu'_{1400} &= N^{(2)}\pi_i\pi_j + 7N^{(3)}\pi_i\pi_j^2 + 6N^{(4)}\pi_i\pi_j^3 + N^{(5)}\pi_i\pi_j^4, \\
\nu'_{2200} &= N^{(2)}\pi_i\pi_j + N^{(3)}\pi_i\pi_j(\pi_i+\pi_j) + N^{(4)}\pi_i^2\pi_j^2, \\
\nu'_{2300} &= N^{(2)}\pi_i\pi_j + N^{(3)}\pi_i\pi_j(\pi_i+3\pi_j) + N^{(4)}\pi_i\pi_j^2(3\pi_i+\pi_j) + N^{(5)}\pi_i^2\pi_j^3, \\
\nu'_{2400} &= N^{(2)}\pi_i\pi_j + N^{(3)}\pi_i\pi_j(\pi_i+7\pi_j) + N^{(4)}\pi_i\pi_j^2(7\pi_i+6\pi_j) + N^{(5)}\pi_i\pi_j^3(6\pi_i+\pi_j) + N^{(6)}\pi_i^2\pi_j^4, \\
\nu'_{1110} &= N^{(3)}\pi_i\pi_j\pi_k, \\
\nu'_{1210} &= N^{(3)}\pi_i\pi_j\pi_k + N^{(4)}\pi_i\pi_j^2\pi_k, \\
\nu'_{1220} &= N^{(3)}\pi_i\pi_j\pi_k + N^{(4)}\pi_i\pi_j\pi_k(\pi_j+\pi_k) + N^{(5)}\pi_i\pi_j^2\pi_k^2, \\
\nu'_{2220} &= N^{(3)}\pi_i\pi_j\pi_k + N^{(4)}\pi_i\pi_j\pi_k(\pi_i+\pi_j+\pi_k) + N^{(5)}\pi_i\pi_j\pi_k(\pi_i\pi_j+\pi_i\pi_k+\pi_j\pi_k) + N^{(6)}\pi_i^2\pi_j^2\pi_k^2, \\
\nu'_{1111} &= N^{(4)}\pi_i\pi_j\pi_k\pi_l.
\end{aligned}\right\} \tag{27}$$

The remaining moments are easily obtained by permutations of k_1, k_2, k_3, k_4 and the corresponding i, j, k, l.

The central multivariate-moments $\nu_{k_1 k_2 k_3 \ldots k_n}$ can be obtained from the symbolic identity

$$\nu'_{k_1 k_2 k_3 \ldots k_n} = (\nu + N\pi_{l_1})^{k_1}(\nu + N\pi_{l_2})^{k_2}(\nu + N\pi_{l_3})^{k_3}\ldots(\nu + N\pi_{l_n})^{k_n}, \tag{28}$$

this being a generalization of the formula given by Kendall (1945) on p. 79. Finally, we have

$$\mu_{k_1 k_2 k_3 \ldots k_n} = \frac{1}{N^{k_1+k_2+k_3+\ldots+k_n}}\nu_{k_1 k_2 k_3 \ldots k_n}. \tag{29}$$

The transfer of the moments is again connected with a great deal of tedious algebra. The central moments required are:

$$
\begin{aligned}
&\mu_{1000} = 0,\\
&\mu_{2000} = \frac{1}{N}\pi_i(1-\pi_i),\\
&\mu_{3000} = \frac{1}{N^2}\pi_i(1-3\pi_i+2\pi_i^2),\\
&\mu_{4000} = \frac{3}{N^2}\pi_i^2(1-2\pi_i+\pi_i^2)+\frac{1}{N^3}\pi_i(1-7\pi_i+12\pi_i^2-6\pi_i^3),\\
&\mu_{5000} = \frac{10}{N^3}\pi_i^2(1-4\pi_i+5\pi_i^2-2\pi_i^3)+\frac{1}{N^4}\pi_i(1-15\pi_i+50\pi_i^2-60\pi_i^3+24\pi_i^4),\\
&\mu_{6000} = \frac{15}{N^3}\pi_i^2(1-3\pi_i+3\pi_i^2-\pi_i^3)+\frac{5}{N^4}\pi_i^2(5-36\pi_i+83\pi_i^2-78\pi_i^3+26\pi_i^4)\\
&\qquad\qquad +\frac{1}{N^5}\pi_i(1-31\pi_i+180\pi_i^2-390\pi_i^3+360\pi_i^4-120\pi_i^5),\\
&\mu_{1100} = -\frac{1}{N}\pi_i\pi_j,\\
&\mu_{1200} = -\frac{1}{N^2}\pi_i\pi_j(1-2\pi_j),\\
&\mu_{1300} = -\frac{3}{N^2}\pi_i\pi_j^2(1-\pi_j)-\frac{1}{N^3}\pi_i\pi_j(1-6\pi_j+6\pi_j^2),\\
&\mu_{1400} = -\frac{10}{N^3}\pi_i\pi_j^2(1-3\pi_j+2\pi_j^2)-\frac{1}{N^4}\pi_i\pi_j(1-14\pi_j+36\pi_j^2-24\pi_j^3),\\
&\mu_{2200} = \frac{1}{N^2}\pi_i\pi_j(1-\pi_i-\pi_j+3\pi_i\pi_j)-\frac{1}{N^3}\pi_i\pi_j(1-2\pi_i-2\pi_j+6\pi_i\pi_j),\\
&\mu_{2300} = \frac{1}{N^3}\pi_i\pi_j(1-\pi_i-6\pi_j+5\pi_j^2+15\pi_i\pi_j-20\pi_i\pi_j^2)\\
&\qquad\qquad -\frac{1}{N^4}\pi_i\pi_j(1-2\pi_i-6\pi_j+6\pi_j^2+18\pi_i\pi_j-24\pi_i\pi_j^2),\\
&\mu_{2400} = \frac{3}{N^3}\pi_i\pi_j^2(1-\pi_i-2\pi_j+\pi_j^2+6\pi_i\pi_j-5\pi_i\pi_j^2)\\
&\qquad\qquad +\frac{1}{N^4}\pi_i\pi_j(1-\pi_i-17\pi_j+42\pi_j^2-26\pi_j^3+41\pi_i\pi_j-156\pi_i\pi_j^2+130\pi_i\pi_j^3)\\
&\qquad\qquad -\frac{1}{N^5}\pi_i\pi_j(1-2\pi_i-14\pi_j+36\pi_j^2-24\pi_j^3+42\pi_i\pi_j-144\pi_i\pi_j^2+120\pi_i\pi_j^3),\\
&\mu_{1110} = \frac{2}{N^2}\pi_i\pi_j\pi_k,\\
&\mu_{1210} = -\frac{1}{N^2}\pi_i\pi_j\pi_k(1-3\pi_j)+\frac{2}{N^3}\pi_i\pi_j\pi_k(1-3\pi_j),\\
&\mu_{1220} = -\frac{1}{N^3}\pi_i\pi_j\pi_k(2-5\pi_j-5\pi_k+20\pi_j\pi_k)+\frac{2}{N^4}\pi_i\pi_j\pi_k(1-3\pi_j-3\pi_k+12\pi_j\pi_k),\\
&\mu_{2220} = \frac{1}{N^3}\pi_i\pi_j\pi_k(1-\pi_i-\pi_j-\pi_k+3\pi_i\pi_j+3\pi_i\pi_k+3\pi_j\pi_k-15\pi_i\pi_j\pi_k)\\
&\qquad -\frac{1}{N^4}\pi_i\pi_j\pi_k(3-7\pi_i-7\pi_j-7\pi_k+26\pi_i\pi_j+26\pi_i\pi_k+26\pi_j\pi_k-130\pi_i\pi_j\pi_k)\\
&\qquad +\frac{2}{N^5}\pi_i\pi_j\pi_k(1-3\pi_i-3\pi_j-3\pi_k+12\pi_i\pi_j+12\pi_i\pi_k+12\pi_j\pi_k-60\pi_i\pi_j\pi_k),\\
&\mu_{1111} = \frac{3}{N^2}\pi_i\pi_j\pi_k\pi_l-\frac{6}{N^3}\pi_i\pi_j\pi_k\pi_l.
\end{aligned}
\qquad (30)
$$

The other central moments may again be obtained by permutations. We are now in the position to find $\mu_2(o_2)$ by substituting appropriate moments from (30) into (24), leading to

$$\mu_2(o_2) = \frac{4}{N}(\omega_3 - \omega_2^3) + \frac{2}{N^2}(\omega_2 - 6\omega_3 + 5\omega_2^3) - \frac{2}{N^3}(\omega_2 - 4\omega_3 + 3\omega_2^3), \tag{31}$$

the first term of which is the same as equation (10) for $n = 2$. Similarly, $\mu_3(o_2)$ can be found by using equations (23) and (30), i.e.

$$\left.\begin{aligned}\mu_3(o_2) = {} & \frac{8}{N^2}(4\omega_4 + 5\omega_2^3 - 9\omega_2\omega_3) - \frac{8}{N^3}(24\omega_4 - 4\omega_3 + 3\omega_2^2 + 22\omega_2^3 - 45\omega_2\omega_3) \\ & + \frac{4}{N^4}(88\omega_4 - 24\omega_3 + \omega_2 + 64\omega_2^3 + 15\omega_2^2 - 144\omega_2\omega_3) \\ & - \frac{4}{N^5}(48\omega_4 - 16\omega_3 + \omega_2 + 30\omega_2^3 + 9\omega_2^2 - 72\omega_2\omega_3),\end{aligned}\right\} \tag{32}$$

and to order N^{-2}

$$\mu_4(o_2) = \frac{48}{N^2}(\omega_3 - \omega_2^2)^2 + O(N^{-3}). \tag{33}$$

By virtue of equations (31)–(33):

$$\beta_1(o_2) = \frac{(4\omega_4 + 5\omega_2^3 - 9\omega_2\omega_3)^2}{(\omega_3 - \omega_2^2)^3 N} + O(N^{-2}), \quad \beta_2(o_2) = 3 + O(N^{-1}), \tag{34}$$

and for large N

$$\beta_1(o_2) \doteqdot 0, \quad \beta_2(o_2) \doteqdot 3,$$

which indicates the convergence of the sampling distribution of o_2 towards normality.

For the special law

$$y\,dx = \frac{1}{10\sqrt{(2\pi)}} e^{-\frac{1}{2}(x/10)^2}\,dx,$$

for $r = 15$ (classes) of width $\frac{1}{2}\sigma = 5$ taken symmetrically from $x = 0$ we have $\omega_2 = 0{\cdot}139608$ if we consider the class width as unity in comparison to $\Omega_2 = 0{\cdot}141047$ for $r = \infty$. Further,

$$\omega_3 = 0{\cdot}022506, \quad \omega_4 = 0{\cdot}003848.$$

A comparison is given below between the first three exact moments of o_2 and their approximations by the first terms in equations (22), (31) and (32) for a sample of $N = 1000$.

Moments	Exact	First-term approximation
$\mu_1'(o_2)$	0·1396	0·1405
$\mu_2(o_2)$	$0{\cdot}1227 \times 10^{-4}$	$0{\cdot}1206 \times 10^{-4}$
$\mu_3(o_2)$	$0{\cdot}5921 \times 10^{-8}$	$0{\cdot}5756 \times 10^{-8}$
$\beta_1(o_2)$	0·0190	0·0189

The approximation (34) to the exact value of $\beta_1(o_2)$ is apparently satisfactory for $N > 1000$. In view of the above, $\beta_2(o_2)$ cannot be substantially different from 3 for $N > 1000$. Hence it may be concluded that for $N > 1000$ the statistic o_2 is reasonably normally distributed. $o_{\frac{1}{2}}$, the other statistic used in practical curve fitting, may also be expected to follow the normal law for $N > 1000$, as it is of lower order than o_2. The same argument, however, will not hold in the case of $a_{\frac{1}{2}}$. Analogous to the moment ratio b_2, its sampling distribution will probably turn out to be skew even for large sample sizes.

Estimation of the parameters of a Pearson Type VII population from large samples

Fisher (1921) has shown that the estimation of the parameters of a Pearson Type VII population

$$f(x)\,dx = \frac{\Gamma(m)}{\sqrt{\pi}\,\Gamma(m-\frac{1}{2})}\frac{c^{2m-1}}{[c^2+(x-\xi)^2]^m}dx \tag{35}$$

is inefficient in the case of $m \leqslant 10$ if carried out by the method of moments. On the other hand, the efficient estimation of the parameters by the method of maximum likelihood is so cumbersome, even using the modified procedure suggested by Jeffreys (1938), that it seems hardly likely that it will ever be adopted by the practical research worker. In the following sections an attempt has been made to show that the method of frequency-moments strikes a balance between the need for efficient estimation and the practical aspect of the method.

Taking the parameters of (35) in the order ξ, c and m we find from maximum likelihood theory the Hessian determinant

$$\Delta = \begin{vmatrix} \frac{(2m-1)\,m}{(m+1)\,c^2} & 0 & 0 \\ 0 & \frac{2m-1}{(m+1)\,c^2} & -\frac{1}{cm} \\ 0 & -\frac{1}{cm} & \left[\frac{d^2\log\Gamma(m-\frac{1}{2})}{dm^2} - \frac{d^2\log\Gamma(m)}{dm^2}\right] \end{vmatrix} \tag{36}$$

(for definition of Hessian see Kendall, 1945, vol. 2, p. 36). From the various zero cross-terms it can be seen that $\hat{\xi}$ is uncorrelated with both $\hat{c}$ and $\hat{m}$, where $\hat{\xi}$, $\hat{c}$ and $\hat{m}$ are maximum likelihood estimates of ξ, c and m. For large N we have, therefore, for either single or joint estimation

$$\text{var}\,(\hat{\xi}) = \frac{(m+1)}{(2m-1)\,m}\frac{c^2}{N}. \tag{37}$$

For estimation of c when m is known and for estimation of m when c is known

$$\text{var}\,(\hat{c}) = \frac{m+1}{2m-1}\frac{c^2}{N}, \tag{38}$$

$$\text{var}\,(\hat{m}) = \frac{1}{[F(m-\frac{3}{2}) - F(m-1)]\,N}, \tag{39}$$

writing

$$F(m) = \frac{d^2\log\Gamma(m+1)}{dm^2}.$$

For the joint estimation of c and m the variances of the maximum likelihood estimators are given by

$$\begin{aligned}\text{var}\,(\hat{c}_J) &= \frac{1}{N\Delta}[F(m-\tfrac{3}{2}) - F(m-1)] \\ &= \frac{m^2(m+1)\,[F(m-\frac{3}{2}) - F(m-1)]}{m^2(2m-1)\,[F(m-\frac{3}{2}) - F(m-1)] - (m+1)} \times \frac{c^2}{N}\end{aligned} \tag{40}$$

and

$$\begin{aligned}\text{var}\,(\hat{m}_J) &= \frac{1}{N\Delta} \times \frac{2m-1}{(m+1)\,c^2} \\ &= \frac{m^2(2m-1)}{m^2(2m-1)\,[F(m-\frac{3}{2}) - F(m-1)] - (m+1)} \times \frac{1}{N}.\end{aligned} \tag{41}$$

All of these results, with the exception of (40), were derived in Fisher's original paper.

Denoting by a bar the estimators based on ordinary moments, we have

$$\text{var}(\bar{\xi}) = \frac{c^2}{(2m-3)N}, \tag{42}$$

and for estimation of c and m when the other one is known

$$\text{var}(\bar{c}) = \frac{m-1}{2m-5} \times \frac{c^2}{N}, \tag{43}$$

$$\text{var}(\bar{m}) = \frac{(2m-3)^2(m-1)}{(2m-5)N} \tag{44}$$

For the joint estimation of c and m by the method of moments

$$\text{var}(\bar{c}_J) = \frac{(m-1)(2m-3)(8m^3-48m^2+108m-83)}{3(2m-5)(2m-7)(2m-9)} \times \frac{c^2}{N}, \tag{45}$$

$$\text{var}(\bar{m}_J) = \frac{2(m-1)(2m-5)(2m-3)^2(2m^2-5m+12)}{3(2m-7)(2m-9)N}, \tag{46}$$

all of which were derived previously by Fisher (1921), with the exception of (44) and (45).

In practice we only need to consider the following cases:

(a) Estimation of ξ,

(b) Estimation of c when m is known,

(c) Estimation of c and m jointly.

In almost every application we will have to deal with case (a).

(b) arises in the case of estimating the scale parameter of an error distribution, especially if $N < 500$. There is strong theoretical and experimental evidence that the errors of truly independent observations are distributed in a Pearson Type VII law with a shape parameter $3 < m < 5$. Jeffreys (1938, 1939, 1948) has pointed out that the estimation of m from the data of a particular experience becomes unreliable if N is less than 500. In that case it is advisable to assume m as being known *a priori*, in the light of previous experimental evidence, with a magnitude of 4, say.

Most frequently, in practice, we do not know the magnitude of any of the parameters. We then have to estimate them from the data simultaneously, provided N is large enough. This is case (c).

The method of frequency-moments lacks a location estimator. We, therefore, have to use either the mean or the median for estimating ξ. Problem (a) resolves into the question which of the two statistics, mean or median, is the more efficient one and for what range. From (37) and (42) we have, taking the maximum likelihood estimator as standard,

$$\text{Eff.}(\bar{\xi}) = \frac{(m+1)(2m-3)}{m(2m-1)}, \tag{47}$$

where $\bar{\xi}$ is the mean statistic. The variance for the median statistic is

$$\text{var}(\breve{\xi}) = \frac{1}{4Nf^2(\xi)} = \frac{c^2\pi}{4N}\left[\frac{\Gamma(m-\frac{1}{2})}{\Gamma(m)}\right]^2, \tag{48}$$

and its efficiency

$$\text{Eff.}(\breve{\xi}) = \frac{4(m+1)}{\pi m(2m-1)}\left[\frac{\Gamma(m)}{\Gamma(m-\frac{1}{2})}\right]^2. \tag{49}$$

Table 1 gives some numerical values for the respective efficiencies. It can be seen that for $m < 3$ the median statistic is the more efficient estimator for ξ. The exact crossing-over of efficiencies takes place at $m = 2{\cdot}840$.

Table 1

m	Eff. ($\bar{\xi}$)	Eff. ($\check{\xi}$)
1	0·0	0·811
1·5	0·0	0·833
2	0·500	0·811
3	0·800	0·769
4	0·893	0·741
5	0·933	0·723
6	0·955	0·710
7	0·967	0·700
8	0·975	0·693
9	0·980	0·687
10	0·984	0·682
∞	1·000	0·637

In the practical applications of the second part of this paper, the median statistic was used for locating the curve whenever $\tilde{m}$, the frequency-moment estimator, was $\leqslant 2{\cdot}840$. In all other cases the mean statistic was employed.

Estimation problem (*b*) leads to the interesting discovery that, when N is large, the maximum likelihood estimator of c can be expressed in terms of simple working probability-moments. For the likelihood solution of (35) we have

$$\frac{\partial}{\partial c}\Sigma \log y = \frac{2m-1}{c}N - 2cm\sum_{i=1}^{N}\frac{1}{c^2+(x_i-\xi)^2} = 0.$$

This may be written as $$\frac{2m-1}{2m} = \frac{1}{N}\left(\frac{c}{y_0}\right)^{1/m}\sum_{i=1}^{N}\left[\frac{y_0 c^{2m-1}}{(c^2+(x_i-\xi)^2)^m}\right]^{1/m}, \tag{50}$$

where $$y_0 = \frac{\Gamma(m)}{\sqrt{\pi}\,\Gamma(m-\frac{1}{2})}.$$

But for large N, $$\sum_{i=1}^{N}[f(x_i)]^{1/m} \doteqdot N\sum_{i=1}^{N}f(x_i)[f(x_i)]^{1/m} = N o_{(m+1)/m}$$

and (50) becomes $$\hat{c} = \left(\frac{2m-1}{2m}\right)^m \frac{\Gamma(m)}{\sqrt{\pi}\,\Gamma(m-\frac{1}{2})} o_{(m+1)/m}^{-m} = \check{c}_{(m+1)/m}. \tag{51}$$

For a population parameter $m = 1$ we have

$$\hat{c} = \frac{1}{2\pi o_2} = \check{c}_2, \tag{52}$$

and for $m = 2$ $$\hat{c} = \frac{9}{8\pi o_{\frac{3}{2}}^2} = \check{c}_{\frac{3}{2}}. \tag{53}$$

Hence it follows that we will obtain maximum efficiency by estimating c from frequency-moments of order 2 and $\frac{3}{2}$ in the case of $m = 1$ and 2 subject to N being large.

The variance of the frequency-moment estimator $\check{c}$ for small deviations and large N, is, from (51),

$$\operatorname{var}(\check{c}) = \left(\frac{\partial \check{c}}{\partial o_{(m+1)/m}}\right)^2 \operatorname{var}(o_{(m+1)/m})$$

$$= \left(\frac{mc}{\omega_{(m+1)/m}}\right)^2 \operatorname{var}(o_{(m+1)/m}),$$

from (10)

$$\operatorname{var}(o_{(m+1)/m}) = \frac{1}{N}\left(\frac{m+1}{m}\right)^2[\omega_{(m+2)/m}^2 - \omega_{(m+1)/m}^2], \quad \operatorname{var}(\check{c}) = \frac{1}{N}c^2(m+1)^2\left[\frac{\omega_{(m+2)/m}}{\omega_{(m+1)/m}} - 1\right], \tag{54}$$

for Type VII law: $$\omega_n \doteqdot (\sqrt{\pi}\, c)^{1-n}\left[\frac{\Gamma(nm-\frac{1}{2})}{\Gamma(nm)}\right]\left[\frac{\Gamma(m)}{\Gamma(m-\frac{1}{2})}\right]^n. \tag{55}$$

Putting $n = (m+2)/m$ and $n = (m+1)/m$ and substituting into (54) gives finally

$$\text{var}\,(\check{c}) = \frac{m+1}{2m-1} \times \frac{c^2}{N}, \tag{56}$$

which is the same as (38), the variance of the maximum likelihood estimator.

When fitting by frequency-moments in practice, the only sets of statistics used so far were $\check{J}_1$, $\check{J}_2$, and $\check{J}_3$ and $\check{J}_1$, $\check{J}_{\frac{1}{2}}$ and $\check{J}_2$ (Sichel, 1947). It is desirable to keep to this procedure in order to avoid computational complications and, generally, to make the method as simple as possible.

Corresponding to the sets of statistics just mentioned we have for a Type VII law the frequency-moment estimators (52) and (53). For large N their respective variances are from (10)

$$\text{var}\,(\check{c}_2) = \left(\frac{c}{\omega_2}\right)^2 \text{var}\,(o_2) = \frac{4}{N}[\alpha_1 - 1]\, c^2, \tag{57}$$

and

$$\text{var}\,(\check{c}_{\frac{1}{2}}) = \left(\frac{2c}{\omega_{\frac{1}{2}}}\right)^2 \text{var}\,(o_{\frac{1}{2}}) = \frac{9}{N}[\alpha_{\frac{1}{2}} - 1]\, c^2. \tag{58}$$

By use of equation (55) we can express α_1 and $\alpha_{\frac{1}{2}}$ as functions of m. Finally, from (38), (57) and (58),

$$\text{Eff.}\,(\check{c}_2) = \frac{m+1}{2m-1}\Big/4\left[\left(\frac{\Gamma(m-\frac{1}{2})\,\Gamma(3m-\frac{1}{2})}{\Gamma(m)\,\Gamma(3m)}\right)\left(\frac{\Gamma(2m)}{\Gamma(2m-\frac{1}{2})}\right)^2 - 1\right], \tag{59}$$

and

$$\text{Eff.}\,(\check{c}_{\frac{1}{2}}) = \frac{m+1}{2m-1}\Big/9\left[\left(\frac{\Gamma(m-\frac{1}{2})\,\Gamma(2m-\frac{1}{2})}{\Gamma(m)\,\Gamma(2m)}\right)\left(\frac{\Gamma(\frac{3}{2}m)}{\Gamma(\frac{3}{2}m-\frac{1}{2})}\right)^2 - 1\right]. \tag{60}$$

The corresponding efficiency for the solution by moments from equations (38) and (43) is

$$\text{Eff.}\,(\bar{c}) = \frac{(m+1)\,(2m-5)}{(m-1)\,(2m-1)}. \tag{61}$$

Numerical values for the efficiencies (59), (60) and (61) are tabulated in Table 2. From this table we may draw some interesting conclusions:

Table 2

m	Eff. $(\check{c}_2)$	Eff. $(\check{c}_{\frac{1}{2}})$	Eff. $(\bar{c})$
0·5	0·0	0·0	0·0
0·8	0·993	0·904	0·0
1	1·000	0·951	0·0
2	0·962	1·000	0·0
3	0·926	0·992	0·400
4	0·903	0·981	0·714
5	0·887	0·972	0·833
6	0·875	0·965	0·891
7	0·867	0·960	0·923
8	0·860	0·955	0·943
9	0·855	0·952	0·956
10	0·851	0·949	0·965
∞	0·808	0·916	1·000

(1) For strongly leptokurtic populations of the Type VII law (low values of m) the frequency-moment estimators are substantially more efficient than the conventional moment estimators.

(2) $\check{c}_{\frac{1}{2}}$ is a more efficient estimator than $\check{c}_2$, having the remarkable property of varying in efficiency only between 0·92 and 1·00 for such a wide range as $1 \leqslant m \leqslant \infty$.

It is the Type VII law of low m which is of practical importance. For example, a normal curve can very well represent a Type VII law even down to $m = 7$. For a sample size of say $N = 1000$ the χ^2-test would still indicate a very good agreement with normal theory even if the true population follows a Type VII law. For a sample size of $N = 1000$ we shall detect real leptokurtosis only if $m < 7$. It is precisely in this range that the method of frequency-moments scores over the method of moments. It also compares very favourably with the maximum likelihood method because the efficiency of $\check{c}_{\frac{1}{2}}$ varies between 0·95 and 1·00 for $1 \leqslant m \leqslant 7$, the computational work, however, being less laborious.

Finally, we have to consider the most frequent estimation problem (*c*) when c and m are unknown. We have

$$a_{\frac{1}{2}} = \phi(\check{m}_J).$$

For small deviations of $\check{m}$ from m and for large N

$$\text{var}\,(\check{m}_J) = \left(\frac{\partial \alpha_{\frac{1}{2}}}{\partial m}\right)^{-2} \text{var}\,(a_{\frac{1}{2}}) = \left[\alpha_{\frac{1}{2}} \frac{\partial \log \alpha_{\frac{1}{2}}}{\partial m}\right]^{-2} \text{var}\,(a_{\frac{1}{2}}). \tag{62}$$

var $(a_{\frac{1}{2}})$ may be expressed in terms of m with the help of equation (55). (62) then becomes

$$\text{var}\,(\check{m}_J) = \frac{\dfrac{\Gamma(m-\frac{1}{2})}{\Gamma(m)}\left\{\dfrac{4\Gamma(3m-\frac{1}{2})}{\Gamma(3m)}\left[\dfrac{\Gamma(2m)}{\Gamma(2m-\frac{1}{2})}\right]^2 + \dfrac{9\Gamma(2m-\frac{1}{2})}{\Gamma(2m)}\left[\dfrac{\Gamma(\frac{3}{2}m)}{\Gamma(\frac{3}{2}m-\frac{1}{2})}\right]^2 - \dfrac{12\Gamma(\frac{5}{2}m-\frac{1}{2})\,\Gamma(\frac{3}{2}m)\,\Gamma(2m)}{\Gamma(\frac{5}{2}m)\,\Gamma(\frac{3}{2}m-\frac{1}{2})\,\Gamma(2m-\frac{1}{2})}\right\} - 1}{N\{[F(m-1)-F(m-\frac{3}{2})]+2[F(2m-1)-F(2m-\frac{3}{2})]-3[F(\frac{3}{2}m-1)-F(\frac{3}{2}m-\frac{3}{2})]\}^2}$$

The efficiency of the joint estimator $\check{m}_J$ (frequency-moment method) is

$$\text{Eff.}\,(\check{m}_J) = \frac{\text{var}\,(\hat{m}_J)}{\text{var}\,(\check{m}_J)} = \frac{\text{equation (41)}}{\text{equation (63)}}. \tag{64}$$

For the corresponding moment method

$$\text{Eff.}\,(\overline{m}_J) = \frac{\text{var}\,(\hat{m}_J)}{\text{var}\,(\overline{m}_J)} = \frac{\text{equation (41)}}{\text{equation (46)}}. \tag{65}$$

The efficiencies (64) and (65) have been tabulated in Table 3.

Table 3

m	Eff. $(\check{m}_J)$	Eff. $(\overline{m}_J)$
0·5	0·0	0·0
0·7	0·677	0·0
1	0·758	0·0
2	0·697	0·0
3	0·651	0·0
4	0·624	0·0
5	0·607	0·169
6	0·596	0·429
7	0·588	0·594
8	0·582	0·699
9	0·577	0·769
10	0·573	0·818
∞	0·537	1·000

Again, the frequency-moment estimates are much better than the moment estimates in the important range $1 \leqslant m \leqslant 7$. There is a loss of information in comparison to the maximum likelihood solution. However, it is not as serious as it first may seem, because we deal with large samples, so that the variances will be comparatively small. Furthermore, given a particular sample after the sampling has taken place, it is the ratio of the standard errors rather than the ratio of the variances which is of practical importance. This value $\sqrt{[\text{Eff.}\,(\check{m}_J)]}$ is quite reasonable for the range $1 \leqslant m \leqslant 7$. On the other hand, the computational saving is considerable when using the new method.

It is hoped that the frequency-moment method will appeal to the practical research worker as a reasonably efficient method which involves fairly simple arithmetic.

PART 2. PRACTICAL ILLUSTRATIONS

It has been pointed out repeatedly that the method of moments will completely break down when, for a Type VII law, the shape parameter $m \leqslant \frac{5}{2}$. Jeffreys (1948) remarks: 'In this case the expectation of the fourth moment is infinite; but the actual fourth moment of any finite set of observations is necessarily finite and any set of observations derived from such a law would be interpreted as having $m > \frac{5}{2}$.' Again, Fisher (1921) proved that the estimation of the parameters of a Type VII law by the method of moments becomes very inefficient except in the region near normality.

Here, then, is a field for the practical application of the method of frequency-moments because all probability-moments exist and because the efficiencies of estimating the parameters by this method are greater than in the case of the ordinary moment solution.

In testing the goodness of fit we know that the contributions to the total χ^2 are made up of

(*a*) a portion deriving from 'errors of estimation';

(*b*) a portion deriving from 'discrepancies of observations from hypothesis' (Fisher, 1948).

The errors of estimation are dependent on the efficiency of the estimation process. In general, we should expect smaller χ^2's when fitting a Type VII law with $m < 7$ by the frequency-moment method than when fitting by the method of moments. This hypothesis was put to a practical test by graduating observed distributions first by the method of moments and then by the method of frequency-moments.

Unless we employ fine grouping we introduce a bias into the estimation of parameters by equating the nth *working* frequency-moment statistic to the integral of the nth power of the probability law, i.e.

$$h^{1-n} \sum_{i=0}^{r+1} f_i^n \doteqdot N^n \int_{-\infty}^{+\infty} y^n dx = J_n,$$

where h = width of class interval. This difficulty may be overcome by estimating the midordinates u_i of the frequency groups by Hardy's formula (1909)

$$u_i = f_i - \tfrac{1}{24}\Delta^2 f_{i-1},$$

where f_i represent the original frequencies. The sum of the nth powers of the midordinates is an estimate of J_n provided there exists high contact at both ends of the experience.

An example will clearly illustrate the advantage of this procedure.

The frequencies of

$$y\,dx = \frac{3000}{\sqrt{\pi}} e^{-0{\cdot}09x^2}\,dx$$

Table 4

Class interval	Frequency
0– 1	1643
1– 2	1376·5
2– 3	965
3– 4	567
4– 5	279
5– 6	115
6– 7	39·5
7– 8	11·5
8– 9	3
9–10	0·5
10–11	0
Total	5000·0
Total of entire distribution	10000·0

were arranged as shown in Table 4. Using uncorrected frequencies we have

$$J_1 = 10000, \quad J_2 = 11879366, \quad \bar{\rho}_1 = \frac{J_2}{J_1} = 8{\cdot}41796.$$

In a normal population $\sigma = 0{\cdot}282095\rho_1$. Hence $s = 2{\cdot}3747$ as compared with $\sigma = 2{\cdot}3570$.

For the corrected frequencies (by Hardy's rule) we find

$$\check{J}_1 = 10000, \quad \check{J}_2 = 11965747, \quad \check{\rho}_1 = 8{\cdot}35719.$$

Hence $s_{\text{cor.}} = 2{\cdot}3575$, which is very near to the true value.

For the purpose of testing the effect of grouping, two adjacent frequency groups were amalgamated and finally three groups, each set giving rise to two possible variations. When applying the midordinate rule it sometimes happens that there is a negative adjustment to a cell at the end of the experience which in the original series had no frequency. In such a case it is best to combine two or three groups and apply their total adjustments jointly.

Table 5 gives a comparison of corrected and uncorrected standard deviations as estimated from the various groupings of the original frequencies of Table 4. The adjustments are not as effective as Sheppard's corrections are to the raw moments, but in all cases they decrease the bias considerably.

Table 5

Class interval	Uncorrected s	Bias $s-\sigma$	Corrected $s_{\text{cor.}}$	Bias $s_{\text{cor.}}-\sigma$
1	2·375	0·018	2·358	0·001
2	2·427	0·070	2·365	0·008
2	2·427	0·070	2·365	0·008
3	2·518	0·161	2·396	0·039
3	2·512	0·155	2·389	0·032

When fitting a Type VII law distribution by the method of frequency-moments we must solve the equation

$$a_{\frac{1}{2}} = \frac{\Gamma(\check{m}-\frac{1}{2})\,\Gamma(2\check{m}-\frac{1}{2})}{\Gamma(\check{m})\,\Gamma(2\check{m})}\left[\frac{\Gamma(\frac{3}{2}\check{m})}{\Gamma(\frac{3}{2}\check{m}-\frac{1}{2})}\right]^2$$

for $\check{m}$. Table 6 facilitates the computations to a great extent. The slight irregularities in the δ^2 column are due to the uncertainty of the last digit in the calculation of $a_{\frac{1}{2}}$. No attempt of adjusting the second differences was made.

Table 6

$\check{m}$	$a_{\frac{1}{2}}$	δ^2+	δ^4+	$\check{m}$	$a_{\frac{1}{2}}$	δ^2+	$\check{m}$	$a_{\frac{1}{2}}$	δ^2+
0·5	∞			3·7	1·08289	6	9·0	1·06869	5
0·6	2·09618			3·8	1·08218	4	9·5	1·06823	5
0·7	1·54561	37382		3·9	1·08151	4	10·0	1·06782	4
0·8	1·36886	9132	22688	4·0	1·08088	3	10·5	1·06745	4
0·9	1·28343	3570	3737	4·1	1·08028	4	11·0	1·06712	3
1·0	1·23370	1745	1056	4·2	1·07972	3	11·5	1·06682	3
1·1	1·20142	976	391	4·3	1·07919	3	12·0	1·06655	2
1·2	1·17890	598	172	4·4	1·07869	3	12·5	1·06630	2
1·3	1·16236	392	83				13·0	1·06607	2
1·4	1·14974	269	48	4·4	1·07869	11	13·5	1·06586	1
1·5	1·13981	194	22	4·6	1·07777	8	14·0	1·06566	2
1·6	1·13182	141	21	4·8	1·07693	9	14·5	1·06548	1
1·7	1·12524	109		5·0	1·07618	6	15·0	1·06531	1
1·8	1·11975	84		5·2	1·07549	5			
1·9	1·11510	66		5·4	1·07485	6	15·0	1·06531	5
2·0	1·11111	53		5·6	1·07427	5	16·0	1·06501	3
2·1	1·10765	44		5·8	1·07374	3	17·0	1·06474	3
2·2	1·10463	36		6·0	1·07324	5	18·0	1·06450	3
2·3	1·10197	29		6·2	1·07279	2	19·0	1·06429	2
2·4	1·09960	26		6·4	1·07236	3	20·0	1·06410	2
2·5	1·09749	21		6·6	1·07196	3	21·0	1·06393	2
2·6	1·09559	19		6·8	1·07159	2	22·0	1·06378	
2·7	1·09388	15		7·0	1·07124	3			
2·8	1·09232	14		7·2	1·07092	1	∞	1·06066	
2·9	1·09090	12		7·4	1·07061	2			
3·0	1·08960	11		7·6	1·07032	2			
3·1	1·08841	9		7·8	1·07005	1			
3·2	1·08731	8		8·0	1·06979	1			
3·3	1·08629	8		8·2	1·06954	2			
3·4	1·08535	6		8·4	1·06931	1			
3·5	1·08447	7		8·6	1·06909	2			
3·6	1·08366	4		8·8	1·06889	0			
3·7	1·08289	6		9·0	1·06869	1			

A detailed calculation of the various statistics involved in the fitting of a Type VII distribution by the method of frequency-moments is shown in Table 7. The observed experience is one of the marginal distributions of the bivariate table, given in Shewhart's book (1931, p. 402), relating to the distribution of random noises (machine measures).

Now
$$\check{J}_n = \Sigma(u_i)^n, \quad o_{\frac{1}{2}} = \frac{\check{J}_{\frac{1}{2}}}{\check{J}_1^{\frac{1}{2}}} = 0{\cdot}38603, \quad a_{\frac{1}{2}} = \frac{o_2}{o_1^2} = \frac{\check{J}_1\check{J}_2}{\check{J}_1^2} = 1{\cdot}08387.$$

From Table 6 we see that for this value of $a_{\frac{1}{2}}$

$$3{\cdot}5 < \check{m} < 3{\cdot}6,$$

and by inverse interpolation we find $\check{m} = 3{\cdot}574$.

Table 7

Central value	f_i	Δf_i	$\Delta^2 f_{i-1}$	$-\frac{1}{24}\Delta^2 f_{i-1}$	u_i	$(u_i)^1$	$(u_i)^2$	f'_i	$f_i - f'_i$	$\frac{(f_i - f'_i)^2}{f'_i}$
0·1349	0							0·9		
0·1421	0	0	1	0				0·7		
0·1493	1	1	−1	0	1·0	1·00	1·00	1·4	− 3·7	0·999
0·1565	1	0	7	−0·3	0·7	0·59	0·49	3·2		
0·1637	8	7	8	−0·3	7·7	21·33	59·29	7·5		
0·1709	23	15	9	−0·4	22·6	107·35	510·76	18·6	+ 4·4	1·041
0·1781	47	24	26	−1·1	45·9	310·74	2106·81	45·5	+ 1·5	0·049
0·1853	97	50	45	−1·9	95·1	927·22	9044·01	101·7	− 4·7	0·217
0·1925	192	95	−62	+2·6	194·6	2714·67	37869·16	181·8	+10·2	0·572
0·1997	225	33	−86	+3·6	228·6	3456·43	52257·96	224·4	+ 0·6	0·002
0·2069	172	−53	−22	+0·9	172·9	2273·63	29894·41	180·0	− 8·0	0·356
0·2141	97	−75	23	−1·0	96·0	940·80	9216·00	100·0	− 3·0	0·090
0·2213	45	−52	26	−1·1	43·9	291·06	1927·21	44·6	+ 0·4	0·004
0·2285	19	−26	14	−0·6	18·4	78·94	338·56	18·2	+ 0·8	0·035
0·2357	7	−12	12	−0·5	6·5	16·57	42·25	7·4		
0·2429	7	0	− 6	+0·3	7·3	19·71	53·29	3·1		
0·2501	1	− 6	5	−0·2	0·8	0·71	0·64	1·4	+ 1·5	0·167
0·2573	0	− 1	1	0·0				0·7		
0·2645	0	0						0·9		
Totals	942	0	0	0·0	942·0	11160·75	143321·84	942·0	0·0	3·532

Further, we have $$\check{c} = \frac{1}{\sqrt{\pi}\, o_1^2}\left[\frac{\Gamma(\check{m})}{\Gamma(\check{m}-\frac{1}{2})}\right]^3\left[\frac{\Gamma(\frac{3}{2}\check{m}-\frac{1}{2})}{\Gamma(\frac{3}{2}\check{m})}\right]^2.$$

Hence $$\check{c} = 3{\cdot}91345.$$

Finally, $$y_0 = \frac{\check{J}_1\,\Gamma(\check{m})}{\check{c}\sqrt{\pi}\,\Gamma(\check{m}-\frac{1}{2})} = 228{\cdot}66,$$

and the equation of the fitted curve taking the class interval as unity

$$y = 228{\cdot}66(1 + 0{\cdot}065295x^2)^{-3{\cdot}574}.$$

As $\check{m} > 2{\cdot}84$ we use the mean statistic for locating the experience. Centring the frequency group 225 at zero on the x-scale and having the positive direction of the scale downwards in the f_i column of Table 7, we find

$$\bar{x} = -0{\cdot}011677.$$

Ordinates were calculated at the beginning, centre and end of each group, and areas were found by a quadrature formula. The expected frequencies are given in the f'_i column of Table 7. For a $\chi^2 = 3{\cdot}532$ we find for 7 degrees of freedom

$$P = 0{\cdot}83.$$

A corresponding moment fit leads to the equation

$$y = 215{\cdot}52(1 + 0{\cdot}029992x^2)^{-6{\cdot}227},$$

with $$\chi^2 = 6{\cdot}756 \quad \text{and} \quad P = 0{\cdot}46.$$

In this particular case the frequency-moment method gives, therefore, a better fit than the conventional method.

Table 8. *Graduation of 15 experiences by a Type VII distribution*

(Column numbers correspond to data derived from 'Sources of Data' listed under the references)

Table 8. *Graduation of 15 experiences by a Type VII distribution*

(Column numbers correspond to data derived from 'Sources of Data' listed under the references)

Central value x	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	Central value z	(12)	(13)	(14)	(15)
−10 and under						6	15	3				+ 0·5	184	580·5	114	283·5
− 9		1				3	3	3				+ 1·5	178	590·5	84	204·5
− 8		1				13	8	0				+ 2·5	154	456·5	53	137·5
− 7		1	1			9	14	3	1	1		+ 3·5	138	354·5	24	78
− 6		6	1			10	34	10	1	6		+ 4·5	122	310	14	63
− 5		19	8		1	31	38	18	4	9		+ 5·5	69	162·5	6	34
− 4	3	23	23	11	6	65	93	26	6	9	3	+ 6·5	60	145·5	3	29
− 3	21	27	47	58	10	147	195	43	14	16	15	+ 7·5	33	98	1	14·5
− 2	78	43	97	115	33	342	349	63	61	32	33	+ 8·5	29	58	1	9
− 1	182	67	192	488	78	622	682	165	194	46	126	+ 9·5	16	44		15
0	305	75	225	933	108	995	990	179	424	65	332	+ 10·5	19	49		20
+ 1	209	42	172	785	103	901	994	88	202	41	100	+ 11·5	9	40		12
+ 2	81	30	97	326	37	702	746	38	59	24	30	+ 12·5	5	37		12·5
+ 3	21	13	45	95	11	359	411	26	23	11	14	+ 13·5	5	19		8·5
+ 4	5	5	19	62	2	170	175	16	3	6	8	+ 14·5	2	26		9·5
+ 5		3	7	12	3	72	98	7	4	7	1	+ 15·5	1	17		27·5
+ 6		3	7			42	57	4	4	0	1	+ 16·5	1	5		12
+ 7		0	1			23	34	5	1	1		+ 17·5	1	5		30
+ 8		1				8	24	2		2		+ 18·5		3		
+ 9						6	24	0		1		+ 19·5		2		
+ 10 and over						14	29	1								
N	905	337	942	2885	392	4540	5013	700	1000	274	670		1026	3003	300	1000
$\frac{m_3}{m}$	5·188	3·261	3·574	3·339	4·546	2·797	2·260	1·448	1·776	2·070	1·432		6·794	2·389	4·005	1·035
$\frac{m_4}{m}$	14·064	5·363	6·227	5·658	5·716	3·272	3·300	3·130	3·230	4·544	3·442		8·384	4·632	6·058	2·726
$\bar{c}^{-2}$	0·08100	0·05806	0·06530	0·14580	0·07766	0·07753	0·09377	0·32135	0·61829	0·11403	1·4356		4·4693 $\times 10^{-2}$	0·02009	0·03783	0·20879
$\bar{c}^{-3}$	0·02474	0·02747	0·02999	0·06638	0·05515	0·05600	0·04299	0·05026	0·17039	0·02798	0·18939		3·4311 $\times 10^{-3}$	0·00652	0·02140	0·00279
$\frac{\bar{\gamma}_1}{\gamma_1}$ %	306·37	72·77	228·66	1002·41	120·21	1024·14	1071·14	191·93	455·35	60·46	384·41		190·36	618·38	118·96	304·85
$\frac{\bar{\gamma}_2}{\gamma_2}$ %	293·07	67·73	215·52	929·64	115·81	964·95	938·59	136·96	367·62	50·42	270·46		187·42	539·65	114·14	84·14
$\frac{\bar{\chi}^2}{\chi^2}$	2·22	6·14	3·53	58·57	3·59	19·68	25·24	21·40	6·23	2·48	6·06		10·27	64·40	0·46	13·55
$\frac{\bar{\chi}^2}{\chi^2}$	3·85	5·73	6·76	77·90	3·90	35·16	120·42	54·90	29·13	8·37	53·28		10·75	117·66	1·02	1016·85
$\frac{\bar{P}}{P}$	0·53	0·42	0·83	< 0·001	0·31	0·08	0·04	0·01	0·29	0·78	0·20		0·42	< 0·001	0·93	0·56
$\frac{\bar{P}}{P}$	0·28	0·46	0·46	< 0·001	0·28	< 0·001	< 0·001	< 0·001	< 0·001	0·14	< 0·001		0·38	< 0·001	0·80	< 0·001

Altogether fifteen observed distributions, which are both symmetric and leptokurtic, were fitted. They are reported in Table 8, and the sources from which they were drawn are given at the end of the paper. In some cases ((4) (6), (7) and (11)) two or more adjacent cells were combined where many groups were given in the original observations. Distribution (15) was reported in unequal class intervals. It is reproduced without alteration, hence the z-scale of Table 8 does not apply to it. For the estimation of the moments and frequency-moments of (15) the coarser class intervals were split arbitrarily so as to make the calculations at all possible. The χ^2-test, however, was applied to the original grouping as shown in column (15).

Distributions (1)–(5) were located by their mean statistics as their $\check{m}$'s all $> 2{\cdot}84$. The median statistics were used for locating distributions (6) to (11) as $\check{m} < 2{\cdot}84$. Experiences (12)–(15) were originally reported without distinction between positive and negative deviations from the mean. It was assumed that their population means are zero. Consequently one further degree of freedom was allowed in testing for goodness of fit.

For lack of space frequencies in some cases ((6), (7) and (8)) falling into cells $+10$ and over and -10 and under were lumped together. In the actual computations, however, the original tail groups were used. The brackets indicate the tail groupings employed in the χ^2-test for both methods of fitting.

In cases where $\check{m} < 3$ it was found necessary to compute more than three ordinates per group, at least in the centre of the distribution, in order to make the error of the quadrature formula negligible.

It is well known that the χ^2-test for goodness of fit is exacting whenever N is very large. In such cases the probability P associated with an observed χ^2 will often be $< 0{\cdot}05$, although the actual fit seems to be quite good (Elderton, 1938, p. 204). For this reason it is better to compare the actual χ^2's as derived from the two methods of fitting instead of their respective probabilities.

A comparison of the χ^2 rows of Table 8 shows that the method of frequency moments yields a better fit in fourteen out of fifteen Type VII distributions examined. The reduction of total χ^2 is substantial in ten out of fifteen cases. As the various cells used were identical for both methods, and as the amalgamation of tail groupings for the χ^2-test were kept the same, it is reasonable to assume that the almost all-round lowering of χ^2's is due to a reduction of errors of estimation. On theoretical grounds we have expected such a result previously.

For distributions (6) and (7) Jeffreys (1939) has estimated the scale and shape parameters of a Type VII law by his modified method of maximum likelihood. A comparison between his method, the frequency and the ordinary moment methods is given below:

Distribution	Statistic \ Method	Modified maximum likelihood	Frequency moments	Moments
(6)	For shape $\hat{m}$	2·710	2·797	3·272
	„ scale $\hat{c}$	3·678	3·591	4·226
(7)	For shape $\hat{m}$	2·257	2·260	3·300
	„ scale $\hat{c}$	3·295	3·266	4·823

Normally, we should not employ the method of frequency moments for the estimation of σ when dealing with a normal universe. There is, however, one exception first pointed out by Yule (1938). We are sometimes confronted with an experience which appears to be reasonably normal except for one or two outlying observations. If the material from which the sample has been drawn is normal, we overestimate the moments due to the disproportionate influence of the outlying observations. On the other hand, it is wrong to omit any observation unless we have evidence of a real blunder in determining or recording an observed quantity.

It is in such a situation that we may prefer the frequency-moment method to the more efficient method of moments (for a normal population) as the former method weights the tail ends far less than the centre of an experience whereas the opposite is true of the moment method.

As an example Pearson's bright line no. 2 experiment (1902) may be quoted. For the moment solution, we find

$$b_1 = 0{\cdot}00178, \quad b_2 = 5{\cdot}02565,$$

which suggests a Type VII law. The fitted equation was found to be

$$y = 100{\cdot}88(1 + 0{\cdot}036624x^2)^{-3{\cdot}981},$$

giving the rather poor fit, $P = 0{\cdot}04$.

Table 9

Central value	f_i	Type VII (Mom.)	Deviations	Normal (Mom.)	Deviations	Normal (Freq. Mom.)	Deviations
+21	1						
+19	0						
+17	0	0·4					
+15	0	0·2	− 4·3		−2·3		− 0·7
+13	0	0·5		0·1		0·1	
+11	0	1·0		0·4		0·2	
+ 9	0	2·0		1·4		1·0	
+ 7	3	4·2		4·4		3·4	
+ 5	8	9·0	− 1·0	11·1	−3·1	9·5	− 1·5
+ 3	31	19·0	+12·0	24·0	+7·0	22·2	+ 8·8
+ 1	35	37·3	− 2·3	43·5	−8·5	42·5	− 7·5
− 1	73	64·6	+ 8·4	65·4	+7·6	66·8	+ 6·2
− 3	76	91·2	−15·2	82·8	−6·8	86·2	−10·2
− 5	96	99·1	− 3·1	87·2	+8·8	91·5	+ 4·5
− 7	79	81·7	− 2·7	77·1	+1·9	79·7	− 0·7
− 9	60	52·9	+ 7·1	56·7	+3·3	57·0	+ 3·0
−11	30	28·8	+ 1·2	35·0	−5·0	33·5	− 3·5
−13	17	14·2	+ 2·8	18·1	−1·1	16·2	+ 0·8
−15	5	6·7		7·8		6·4	
−17	3	3·1		2·8		2·1	
−19	1	1·5		0·9		0·6	
−21	0	0·7	− 2·9	0·2	−1·8	0·1	+ 0·8
−23	0	0·4		0·1			
−25	0	0·2					
−27	1	0·3					
P			0·04		0·44		0·56

For the method of frequency-moments we have

$$a_4 = 1{\cdot}06185.$$

From Table 6 we see that $\check{m}$ must be very far above 22. As the difference between a normal curve and a Type VII curve of $m > 30$ is very minute we should proceed to fit a normal curve by the method of frequency moments. This was done giving

$$y = 92{\cdot}57e^{-0{\cdot}09996x^2},$$

with the good fit $P = 0{\cdot}56$.

Had we known *a priori* that the sample came from a normal population we should have fitted a normal curve by the method of moments leading to the equation

$$y = 88{\cdot}27e^{-0{\cdot}09086x^2}, \quad P = 0{\cdot}44.$$

The original observations and the resulting three fits are given in Table 9.

Inspection of Table 9 shows that the frequency-moment fit is better than the ordinary moment solution as

(1) P is larger;

(2) absolute deviations are smaller in ten out of twelve cells;

(3) there are nine changes of signs as compared with six in the moment fit.

The use of Hardy's formula in the practical applications of the method of frequency-moments introduces certain inconsistencies with regard to the standard errors and efficiencies as derived in the theoretical part of the investigation. It is confidently felt, however, that the discrepancy between the practical and theoretical approach is small, just as, in practice, one works so often with Sheppard's correction without taking it into account when estimating standard errors.

It gives me great pleasure to express my thanks to Mr J. E. Kerrich for much helpful criticism, to the South African Council for Scientific and Industrial Research for permission to publish this paper, and to the staff of the Statistical Section for assisting me in heavy computational work.

REFERENCES

ELDERTON, W. P. (1938). *Frequency Curves and Correlation.* Cambridge University Press.

FISHER, R. A. (1921). On the mathematical foundations of theoretical statistics. *Philos. Trans.* A, **222**, 309.

FISHER, R. A. (1948). *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.

HARDY, G. F. (1909). *The Theory of the Construction of Tables of Mortality and of Similar Statistical Tables in Use by the Actuary.* London: Institute of Actuaries.

JEFFREYS, H. (1938). The law of errors and the combination of observations. *Philos. Trans.* A, **237**, 231.

JEFFREYS, H. (1939). The law of errors in the Greenwich variation of latitude observations. *Mon. Not. R. Astr. Soc.* **99**, 703.

JEFFREYS, H. (1948). *Theory of Probability.* Oxford: Clarendon Press.

KENDALL, M. G. (1945). *The Advanced Theory of Statistics*, **1**. London: Charles Griffin and Co. Ltd.

PEARSON, K. (1902). On the mathematical theory of errors. *Philos. Trans.* A, **198**, 235.

SHEWHART, W. A. (1931). *Economic Control of Quality of Manufactured Product.* New York: van Nostrand.

SICHEL, H. S. (1947). Fitting growth and frequency curves by the method of frequency moments. *J.R. Statist. Soc.* **110**, 337.

YULE, G. U. (1938). On some properties of normal distributions, univariate and bivariate, based on sums of squares of frequencies. *Biometrika*, **30**, 1.

SOURCES OF DATA

(1) TERMAN, L. M. (1928). Intelligence quotients. From *The Measurement of Intelligence*. London: Harrap and Co. Ltd.
(2) NAT. INST. PERS. RES. Visual discrimination scores. Unpublished research data.
(3) SHEWHART, W. A. (1931). Random noises. From *Economic Control of Quality of Manufactured Product*. New York: van Nostrand.
(4) MICHELSON, A. A. and others (1935). Light velocities. *J. Astrophys.* **32**, 56.
(5) FLEMINGER, M. E. Output deviations. Unpublished research data.
(6, 7) HULME, H. R. & SYMMS, L. S. J. (1939). Variation of latitude. *Mon. Not. R. Astr. Soc.* **99**, 642.
(8) HANSMANN, G. H. (1934). Distribution of third moment coefficients. *Biometrika*, **26**, 129.
(9) SICHEL, H. S. (1947). Sampling experiment on Type VII population. *J. Roy. Statist. Soc.* **110**, 337.
(10) KERRICH, J. E. Fortnightly gains of oxen. Unpublished research data.
(11) MILLS, F. C. (1938). Prices and commodities in 1927. From *Statistical Methods*. New York: Holt and Co.
(12) BOND, W. N. (1935). Variation of judgement on the position of fuzzy object. From *Probability and Random Errors*. London: Arnold.
(13) PEARSON, K. & MOUL, M. (1927). Distribution of tetrads. *Biometrika*, **19**, 246.
(14) BESSEL, F. W. (1838). Errors of right ascensions. *Astr. Nachr.* **15**, 358.
(15) PEARSON, E. S. (1929). Student's ratio for samples of two. *Biometrika*, **21**, 259.

ON THE USE OF STUDENT'S t-TEST IN AN ASYMMETRICAL POPULATION

By S. G. GHURYE

On account of the unique property of samples from a normal population that the ratio $(\bar{x}-\mu)\sqrt{(n+1)}/s$ (where μ is the population mean, $\bar{x} = \sum_{i=1}^{n+1} x_i/(n+1)$ and $ns^2 = \Sigma(x-\bar{x})^2$) is the ratio of a normal deviate to a stochastically independent estimate of its variance, Student's t-test is a suitable test of significance for the mean of a normal population. However, in a variety of cases, it is necessary to test for the mean of a population which does not follow the Gaussian law. Efforts have, therefore, been made to see how far Student's distribution may be used for the purpose in non-normal populations. Due, mainly, to the analytical difficulties of the problem, no extensive theoretical discussion has yet been given. Thus, Pearson & Adyanthaya (1929), Rietz (1939) and Nair (1941) have given experimental treatments, while the theoretical discussions of some others (Rider, 1929; Perlo, 1933; Laderman, 1939) have dealt only with trivially small sample sizes. The papers by Bartlett (1935) and Geary (1936, 1947) give results true for any sample size, though they are based on certain assumptions and approximations. The present paper deals with the population considered by Geary in his 1936 paper, subject to the same approximations. The second contribution by Geary (in which is derived the t-distribution in samples from a population which departs more from normality than that considered in the 1936 paper) came to my notice too late to be made use of in the present work; but it is proposed to consider it later on.

Geary (1936) has obtained the distribution of the ratio $(\bar{x}-\mu)\sqrt{(n+1)}/s$ in the case of an asymmetrical population, whose fourth and higher cumulants are zero, by neglecting squares and higher powers of the third cumulant. We know from this how far the probability of an error of the first kind (i.e. the probability of rejecting the null hypothesis when it is true) in such a population differs from that for a normal distribution, provided we may neglect the square of the standardized third cumulant γ_1. Here again, on account of analytical difficulties, it is not possible, except for very small sample sizes, to consider the effect of terms containing higher powers of γ_1. However, we can assume the result derived by Geary to be correct for very small values of γ_1, as also for large sample sizes—but in such cases the deviation from values of the normal theory is practically negligible. Even then, it is of interest to know whether, in using the usual tables of the t-test (based on the normal distribution), we are committing the greater error in the probability of an error of the first kind or in that of an error of the second kind. In the present paper are derived the values of the probability of an error of the second kind (and hence, of the power of the test) when the usual t-tables are used to define the critical region.

It may be mentioned here that this problem is only a special case of a general investigation, on which the writer is engaged, into the effect, on statistical tests, of differences between the actual and the assumed distribution laws of the universe sampled. The solution of these problems is hampered by analytical difficulties in the derivation of the probability laws (and particularly of power functions), and the present case is one of the few in which a mathematical, though only approximate, solution has been found possible.

EXPRESSION FOR THE POWER FUNCTION

Let the variate x have a mean μ_1, standard deviation σ, third cumulant $\kappa_3 = \gamma_1\sigma^3$, and all higher cumulants zero; γ_1 is assumed to be small, squares and higher powers being neglected. The distribution function of x is given by

$$dF(\xi) = \frac{1}{\sqrt{(2\pi)}}\left\{1+\frac{\gamma_1}{6}(\xi^3-3\xi)\right\}\exp(-\tfrac{1}{2}\xi^2)\,d\xi,$$

where $\xi = \dfrac{x-\mu_1}{\sigma}$. Again neglecting γ_1^2 and higher powers, the probability of a sample of $n+1$ independent values $x_1, x_2, \ldots, x_{n+1}$ is given by

$$(2\pi)^{-\frac{1}{2}(n+1)}\left\{1+\frac{\gamma_1}{6}(\Sigma\xi^3-3\Sigma\xi)\right\}\exp(-\tfrac{1}{2}\Sigma\xi^2)\,d\xi_1\ldots d\xi_{n+1}.$$

The part not containing γ_1 is the same as for the normal population. In what follows, we are assuming that the table of the t-test based on the normal distribution is used for the significance level, i.e. for the probability of an error of the first kind. The contribution of the part not containing γ_1 is, therefore, the same as for a sample from a normal population, for which the power of the t-test has been considered by Neyman *et al.* (1935) and again by Neyman & Tokarska (1936). Hence, we shall consider below only the additive part

$$\frac{\gamma_1}{6(2\pi)^{\frac{1}{2}(n+1)}}(\Sigma\xi^3-3\Sigma\xi)\exp(-\tfrac{1}{2}\Sigma\xi^2)\,d\xi_1\ldots d\xi_{n+1}. \tag{1}$$

The joint distribution of $\bar{x}$ and s can be obtained by the substitution used by Geary (1936), and is found to be

$$dF(u,\chi) = \frac{\gamma_1}{2^{\frac{1}{2}(n-2)}\left(\dfrac{n-2}{2}\right)6\sqrt{[2\pi(n+1)]}}\{u^3+3u\chi^2-3(n+1)u\}\chi^{n-1}\exp\{-\tfrac{1}{2}(\chi^2+u^2)\}du\,d\chi, \tag{2}$$

where $$u = \frac{(\bar{x}-\mu_1)\sqrt{(n+1)}}{\sigma} \quad \text{and} \quad \chi^2 = \frac{\Sigma(x_i-\bar{x})^2}{\sigma^2} = \frac{ns^2}{\sigma^2}.$$

For testing the hypothesis $\mu = \mu_0$, the t-test is uniformly most powerful for the class of one-sided alternatives $\mu > \mu_0$ or $\mu < \mu_0$. Thus, to test the hypothesis $\mu = \mu_0$ (or $< \mu_0$) against the set of alternatives $\mu > \mu_0$, we find the value t of Student's ratio such that the probability of exceeding it is a predetermined α (i.e. the value of t in Fisher's tables corresponding to the probability 2α); we then reject the hypothesis $\mu \leqslant \mu_0$ if the ratio $(\bar{x}-\mu_0)\sqrt{(n+1)}/s$ exceeds t, and accept it if it does not. In both these decisions we may be wrong in that we may decide that $\mu > \mu_0$, when, in fact, $\mu \leqslant \mu_0$ (error of the first kind), or we may accept $\mu \leqslant \mu_0$ even when $\mu = \mu_1 > \mu_0$ (error of the second kind). From Geary's paper we can calculate the difference between the actual value of the probability of an error of the first kind and the value α assumed. To obtain the corresponding correction for the probability of an error of the second kind (and hence for the power of the test), we have to integrate the expression (2) over the region of acceptance, i.e. over the domain conditioned by $\dfrac{(\bar{x}-\mu_0)\sqrt{(n+1)}}{s} \leqslant t$,

i.e. by $\quad u \leqslant \dfrac{\chi t}{\sqrt{n}} - \rho_n, \quad$ where $\quad \rho_n = \dfrac{\mu_1-\mu_0}{\sigma}\sqrt{(n+1)}$ and t is a function of α.

Then the probability of an error of the second kind is

$$P_{\mathrm{II}}(\alpha, n, \rho_n) = P^0_{\mathrm{II}}(\alpha, n, \rho_n) - \gamma_1 P^1_{\mathrm{II}}(\alpha, n, \rho_n), \tag{3}$$

where P_{II}^0 is the P_{II} for the normal population and $-\gamma_1 P_{II}^1$ is the contribution of the terms containing only the first power of γ_1. The power function $(1-P_{II})$ is then $(1-P_{II}^0+\gamma_1 P_{II}^1)$, so that $\gamma_1 P_{II}^1$ is the correction to be added to the 'normal-theory' value. Then

$$P_{II}^1 = \frac{-1}{2^{\frac{1}{2}(n-2)}\left(\frac{n-2}{2}\right|\Big)6\sqrt{[2\pi(n+1)]}}\int_0^\infty\int_{-\infty}^{(t\chi/\sqrt{n})-\rho_n}\{u^3+3u\chi^2-3(n+1)u\}\chi^{n-1}$$
$$\exp\{-\tfrac{1}{2}(\chi^2+u^2)\}\,du\,d\chi.$$

Now, the double integral

$$=\int_0^\infty \exp(-\tfrac{1}{2}\chi^2)\,\chi^{n-1}\left[\int_{-\infty}^{(t\chi/\sqrt{n})-\rho_n}\{u^3+3u\chi^2-3(n+1)u\}\exp(-\tfrac{1}{2}u^2)\,du\right]d\chi$$
$$=-\int_0^\infty\{2+u_0^2+3(\chi^2-\overline{n+1})\}\chi^{n-1}\exp\{-\tfrac{1}{2}(\chi^2+u_0^2)\}\,d\chi, \quad \text{where} \quad u_0=\frac{t\chi}{\sqrt{n}}-\rho_n$$
$$=-\exp\left(-\frac{\rho_n^2}{2a^2}\right)\int_0^\infty\{2+u_0^2+3(\chi^2-\overline{n+1})\}\chi^{n-1}\exp\left\{-\frac{1}{2}\left(\chi a-\frac{t\rho_n}{a\sqrt{n}}\right)^2\right\}d\chi$$
$$=-\frac{\exp(-\rho_n^2/2a^2)}{a^{(n+2)}}\int_{-b}^\infty(Aw^2+4bw+B)(w+b)^{n-1}\exp(-\tfrac{1}{2}w^2)\,dw,$$

in which $a^2=1+\dfrac{t^2}{n}$, $b=\dfrac{t\rho_n}{a\sqrt{n}}$, $A=a^2+2$ · and $B=(3a^2-2)\dfrac{\rho_n^2}{a^2}-(3n+1)a^2$.

Hence
$$P_{II}^1=\frac{\exp(-\rho_n^2/2a^2)\,\mathscr{I}_n}{2^{\frac{1}{2}(n-2)}\left(\frac{n-2}{2}\right|\Big)6\sqrt{(n+1)}\,a^{(n+2)}}, \tag{4}$$

where
$$\mathscr{I}_n=\frac{1}{\sqrt{(2\pi)}}\int_{-b}^\infty(Aw^2+4bw+B)(w+b)^{n-1}\exp(-\tfrac{1}{2}w^2)\,dw. \tag{5}$$

Thus, the problem of finding the values of $\gamma_1 P_{II}^1$, the contribution of the first power of γ_1 to the power function, reduces to that of the evaluation of $\mathscr{I}_n$.

Evaluation of P_{II}^1

By expanding the binomial in (5), $\mathscr{I}_n$ can be expressed as a linear function of $n+2$ incomplete Γ-functions. However, the best method of evaluating $\mathscr{I}_n$, at least for the values of n considered below, seems to be that based on a reduction formula.

Let
$$K_n=\frac{1}{\sqrt{(2\pi)}}\int_{-b}^\infty(w+b)^n\exp(-\tfrac{1}{2}w^2)\,dw, \tag{6}$$

so that
$$\mathscr{I}_n=AK_{n+1}-2a^2bK_n+(\rho_n^2-3n-1)a^2K_{n-1}. \tag{7}$$

From (6), we get by partial integration
$$K_n=bK_{n-1}+(n-1)K_{n-2}. \tag{8}$$

Denoting by f the normal probability-density $\dfrac{1}{\sqrt{(2\pi)}}\exp(-\frac{1}{2}b^2)$ and by F the distribution function $\dfrac{1}{\sqrt{(2\pi)}}\displaystyle\int_{-\infty}^b\exp(-\tfrac{1}{2}u^2)\,du$, we have $K_0=F$ and $K_1=bF+f$, from which successive K's can be evaluated by using (8).*

Then, from (7) and (8),
$$\mathscr{I}_n=lK_n+mK_{n-1}, \tag{9}$$

where
$$l=b(2-a^2) \quad \text{and} \quad m=2n+(\rho_n^2-2n-1)a^2.$$

* It is to be noted that $K_n=n!\,Hh_n(-b)/\sqrt{(2\pi)}$, where Hh_n is the function tabulated (for $n\leqslant 21$) in Table XV of the British Association Mathematical Tables, Vol. I. I am obliged to Dr Harold Hotelling for pointing this out to me.

The values of $P_{II}^{1}(\alpha, n, \rho_n)$ were calculated, by way of illustration, for $\alpha = 0{\cdot}05$ (Table 1) and $\alpha = 0{\cdot}01$ (Table 2), $n = 4$, 9 and 19 and all integral values of ρ_n for which the value of P_{II}^{1} exceeded 0·001 in absolute value. Geary (1936) has tabulated the values of his correction to the distribution function for these values of n. The value of $-P_{II}^{1}$ for $\rho_n = 0$ must, of course, agree with that in Geary's table for the same n and the appropriate t. The accuracy of the results is conditioned by the accuracy of the value of t obtained from the probability table, but the effect of subsequent errors due to rounding off has been rendered negligible by retaining six or seven significant figures in the calculations. Under the circumstances, the results given should be correct, except for the effect of rounding-off errors on the fourth place of decimals.

Table 1. P_{II}^{1} *for* $\alpha = 0{\cdot}05$

n \ ρ_n	0	1	2	3	4	5	6
4	−0·0343	−0·0775	−0·0361	+0·0559	+0·0649	+0·0260	+0·0051
9	−0·0297	−0·0628	−0·0046	+0·0597	+0·0340	+0·0062	—
19	−0·0229	−0·0450	+0·0048	+0·0445	+0·0191	+0·0025	—

Table 2. P_{II}^{1} *for* $\alpha = 0{\cdot}01$

n \ ρ_n	0	1	2	3	4
4	−0·0105	−0·0418	−0·0771	−0·0619	+0·0066
9	−0·0115	−0·0509	−0·0753	+0·0012	+0·0731
19	−0·0098	−0·0444	−0·0523	+0·0270	+0·0569

n \ ρ_n	5	6	7	8	9
4	+0·0616	+0·0646	+0·0387	+0·0156	+0·0045
9	+0·0496	+0·0130	+0·0016	—	—
19	+0·0207	+0·0026	—	—	—

EFFECT ON THE POWER FUNCTION

Following Geary's 1936 paper, it has been assumed that the γ_1 of the population in question is sufficiently small for our purpose; but it has not been found possible to define the range of permissible values of γ_1. It is proposed to consider this question, as also that of improving the approximation to the power function, in a further publication.

In order to get an idea of the magnitude of the effect of departure from normality, the values of the power for $\gamma_1 = 0$ (normal distribution) and $\gamma_1 = 0{\cdot}4$, $n = 9$ and 19 and some values of ρ_n are given in Table 3, assuming that the effect of higher powers of γ_1 is negligible. In this respect, it may be mentioned here that on the wider assumption used in Geary's

1947 paper the actual probability of t lying to the left of the lower 2·5 % point of normal theory is 0·041 in a sample of 10 from a Pearsonian population with $\gamma_1 = 0{\cdot}5$ and $\gamma_2 = 0$, whereas, according to the assumption on which the present paper is based, the same probability is 0·035. It has, therefore, been assumed for the present that the approximation used in the present paper also gives a fair estimate of the power function when $\gamma_1 = 0{\cdot}4$.

Table 3. *Comparison of the power when* $\gamma_1 = 0{\cdot}0$ and $\gamma_1 = 0{\cdot}4$

$\alpha = 0{\cdot}05$					$\alpha = 0{\cdot}01$			
$n = 9$		$n = 19$			$n = 9$		$n = 19$	
$\gamma_1 = 0{\cdot}0$	$\gamma_1 = 0{\cdot}4$	$\gamma_1 = 0{\cdot}0$	$\gamma_1 = 0{\cdot}4$	ρ_n	$\gamma_1 = 0{\cdot}0$	$\gamma_1 = 0{\cdot}4$	$\gamma_1 = 0{\cdot}0$	$\gamma_1 = 0{\cdot}4$
0·050	0·038	0·050	0·041	0	0·010	0·005	0·010	0·006
0·236	0·211	0·248	0·230	1	0·071	0·051	0·081	0·063
0·580	0·578	0·612	0·614	2	0·268	0·238	0·320	0·299
0·868	0·892	0·894	0·912	3	0·587	0·587	0·677	0·688
0·979	0·993	0·986	0·994	4	0·853	0·882	0·916	0·939
				5	0·969	0·989	0·989	0·997

The effect of positive skewness is to decrease the power in the region of less power and increase it everywhere else, the opposite being true of negative skewness. It is seen from Table 3 that the change in the power is not such as to affect materially the inference drawn from the test. How far this is so for greater departures from normality, it is not possible to determine by the present method.

Conclusion

It has thus been found possible to obtain an expression, though approximate, for the power of Student's t-test applied to samples from an asymmetrical universe when the critical region has been determined on the erroneous assumption of normality in the parent population.

The results derived are true if the fourth and higher cumulants are zero and the standardized third cumulant sufficiently small; subject to these restrictions, the change in the power of the test is found to be negligible as far as the inference to be drawn is concerned.

My thanks are due to Dr N. R. Tawde of the Royal Institute of Science, Bombay, and Prof. G. S. Priolker of the Wilson College, Bombay, for allowing me to use their calculating machines.

REFERENCES

Bartlett, M. S. (1935). *Proc. Camb. Phil. Soc.* **31**, 223.
Geary, R. C. (1936). *J.R. Statist. Soc. Suppl.* **3**, 178.
Geary, R. C. (1947). *Biometrika*, **34**, 209.
Laderman, J. (1939). *Ann. Math. Statist.* **10**, 376.
Nair, A. N. K. (1941). *Sankhyā*, **5**, 393.
Neyman, J. *et al.* (1935). *J.R. Statist. Soc. Suppl.* **2**, 107.
Neyman, J. & Tokarska, B. (1936). *J. Amer. Statist. Ass.* **31**, 318.
Pearson, E. S. & Adyanthaya, N. K. (1929). *Biometrika*, **21**, 259.
Perlo, V. (1933). *Biometrika*, **25**, 203.
Rider, P. R. (1929). *Biometrika*, **21**, 124.
Rider, P. R. (1931). *Ann. Math. Statist.* **2**, 48.
Rietz, H. L. (1939). *Ann. Math. Statist.* **10**, 265.

TABLES OF SYMMETRIC FUNCTIONS—PART I

By F. N. DAVID and M. G. KENDALL

1. Symmetric functions have important applications in the theory of probability and the theory of sampling. There are four types of symmetric function in use, and we shall also employ a fifth. They are as follows:

(*a*) The monomial symmetric functions typified by

$$(p_1^{\pi_1} p_2^{\pi_2} \dots p_s^{\pi_s}) = \Sigma x_i^{p_1} x_j^{p_1} \dots x_q^{p_2} x_r^{p_2} \dots x_u^{p_s} x_v^{p_s}, \tag{1}$$

where the summation takes place over all suffixes $i, j, \dots, q, r, \dots, u, v$ which are different. The number $\sum_{j=1}^{s} \pi_j = \varpi$, say, is called the *order* of the symmetric function and $\sum_{j=1}^{s} p_j \pi_j = w$, say, is called its *weight*. There are $\binom{n}{\varpi}$ terms in the summation on the right in (1) where n is the number of possible different suffixes.

(*b*) The unitary symmetric functions (1^r), denoted by a_r. We have, from equation (1),

$$a_r = (1^r) = \Sigma x_i x_j \dots x_v, \tag{2}$$

where the subscripts $i, j, \dots, v$ are all different and there are r of them in any one term of the sum. a_r may also be defined by the identity in t

$$T = (t - x_1)(t - x_2) \dots (t - x_n) = t^n - a_1 t^{n-1} + a_2 t^{n-2} - \dots; \tag{3}$$

(*c*) The one-part symmetric functions or power-sums, defined by

$$s_r = (r) = \sum_{i=1}^{n} x_i^r. \tag{4}$$

These also are a special case of (1).

(*d*) The homogeneous product-sums defined by

$$\frac{1}{T} = 1 + h_1 t + \dots + h_r t^r + \dots;$$

we shall not use these functions. They are mentioned for the sake of completeness.

(*e*) The "augmented" functions $[p_1^{\pi_1} \dots p_s^{\pi_s}] = (p_1^{\pi_1} \dots p_s^{\pi_s}) \pi_1! \dots \pi_s!$. (5)

For many statistical purposes these are more convenient than the ordinary monomials.

2. The tables below show the product of power-sums of weight w in terms of the augmented symmetrics and vice versa, up to and including weight 12. Each table is given the number 1 (because further sets of tables are contemplated) followed by its weight, e.g. Table 1.5 relates to weight 5. To express the power-sums in terms of augmented symmetrics the tables are read horizontally up to and including the diagonal, the unit entry in which is shown in *heavy* type. For example, from Table 1.5

$$(2)^2 (1) = [5] + [41] + 2[32] + [2^2 1]. \tag{6}$$

The augmented symmetrics are given in terms of power-sums by reading the tables vertically from the top, again up to and including the diagonal. For example, from Table 1.5

$$[31^2] = 2(5) - 2(4)(1) - (3)(2) + (3)(1)^2. \tag{7}$$

Uses of the tables

3. A minor use of the tables may be noted: The expression of $(1)^r$ in terms of the augmented symmetrics gives the expansion of moments about an arbitrary point in terms of cumulants; e.g. from the bottom line of Table 1.5 we have

$$\mu'_5 = \kappa_5 + 5\kappa_4\kappa_1 + 10\kappa_3\kappa_2 + 10\kappa_3\kappa_1^2 + 15\kappa_2^2\kappa_1 + 10\kappa_2\kappa_1^3 + \kappa_1^5. \tag{8}$$

If, of course, we can accept moments about the mean we ignore any term containing a unit part; e.g.

$$\mu_5 = \kappa_5 + 10\kappa_3\kappa_2. \tag{9}$$

4. In statistical investigations concerning sampling moments we often need the expectations of products of power-sums. These can be written down at sight, for samples from an infinite population, from Table 1 when it is remembered that

$$\mathscr{E}[p_1^{\pi_1}\dots p_s^{\pi_s}] = n^{(\varpi)}(\mu'_{p_1})^{\pi_1}(\mu'_{p_2})^{\pi_2}\dots(\mu'_{p_s})^{\pi_s}, \tag{10}$$

where
$$n^{(\varpi)} = n(n-1)\dots(n-\varpi+1).$$

Thus, from (6),

$$\mathscr{E}\{(2)^2(1)\} = n\mu'_5 + n(n-1)\mu'_4\mu'_1 + 2n(n-1)\mu'_3\mu'_2 + n(n-1)(n-2)\mu'^2_2\mu'_1. \tag{11}$$

5. The same kind of procedure gives the multivariate cumulants in terms of multivariate moments about an arbitrary point. For example, corresponding to (6) we have

$$\mu'(2^21) = \kappa(2^21) + \kappa(2^2)\kappa(1) + 2\kappa(21)\kappa(2) + \{\kappa(2)\}^2\kappa(1). \tag{12}$$

Conversely, from (7),

$$\kappa(31^3) = 2\mu'(3)\{\mu'(1)\}^2 - 2\mu'(31)\mu'(1) - \mu'(3)\mu'(1^3) + \mu'(31^3). \tag{13}$$

Equations of type (13) are particularly useful.

6. Table 1 may also be used to express k-statistics in terms of the one-part symmetrics (r), namely, in terms of the power-sums which are used to calculate them in practice. The rth k-statistic is the symmetric function of the sample values whose expectation is equal to the parent cumulant of order r. Thus, corresponding to the inverse of (8), we have

$$k_5 = \frac{[5]}{n} - \frac{5[41]}{n(n-1)} - \frac{10[32]}{n(n-1)} + \frac{20[31^2]}{n(n-1)(n-2)} + \frac{30[2^21]}{n(n-1)(n-2)} - \frac{60[21^3]}{n(n-1)(n-2)(n-3)} + \frac{24[1^5]}{n(n-1)(n-2)(n-3)(n-4)}. \tag{14}$$

Substituting for the augmented symmetrics by power-sums written as s_r, from Table 1d, and collecting terms, we get

$$k_5 = \frac{1}{n^{(5)}}\{(n^4+5n^3)s_5 - 5(n^3+5n^2)s_4s_1 - 10(n^3-n^2)s_3s_2 + 20(n^2+2n)s_3s_1^2 + 30(n^2-n)s_2^2s_1 - 60ns_2s_1^3 + 24s_1^5\}. \tag{15}$$

7. The tables provide a method of evaluating the sampling cumulants of k-statistics (or of any symmetric functions of the observations) by the direct evaluation of expectations. The alternative combinatorial method (Fisher, 1929; Kendall, 1941) is shorter in performance but needs careful handling in view of the ease with which certain combinations can be overlooked. Kendall (1941) lists most of the cumulants of k-statistics up to order 12 but does not give $\kappa(3^32)$. We will sketch the derivation of this quantity as an illustration of the use of the tables.

By relations derived as in § 6 we have

$$k_3^3 k_2 = \left[\frac{n}{(n-1)(n-2)}\left\{s_3 - \frac{3s_1s_2}{n} + \frac{2s_1^3}{n^2}\right\}\right]^3 \frac{1}{n-1}\left(s_2 - \frac{s_1^2}{n}\right)$$

$$= \frac{1}{n^4(n-1)^4(n-2)^3}\{n^7 s_3^3 s_2 - n^6(s_3^3 s_1^2 + 9s_3^2 s_2^2 s_1)$$

$$+ n^5(15s_3^2 s_2 s_1^3 + 27s_3 s_2^3 s_1^2) - n^4(6s_3^2 s_1^5 + 63s_3 s_2^2 s_1^4 + 27s_2^4 s_1^3)$$

$$+ n^3(48s_3 s_2 s_1^6 + 81s_2^3 s_1^5) - n^2(12s_3 s_1^8 + 90s_2^2 s_1^7) + n(44s_2 s_1^9) - 8s_1^{11}\}. \quad (16)$$

We take expectations of both sides. The terms on the right may be written down in terms of parent μ's after the manner of § 4. These μ's are then transformed to parent κ's by the known relations connecting them, for example, those of § 3. The result gives us $\mu'(3^32)$ and the corresponding κ is then derived from the equation

$$\kappa(3^32) = \mu'(3^32) - 3\kappa_3\kappa(3^32) - \kappa_2\kappa(3^3) - 3\kappa(3^2)\kappa(32) - 3\kappa(32)\kappa_3^2 - 3\kappa_2\kappa_3\kappa(3^2) - \kappa_3^3\kappa_2, \quad (17)$$

which is itself derived from the tables as described in § 5. The κ's of lower order occurring in this expression have already been tabulated, and we find, after carrying out the necessary substitutions and reductions

$$\kappa(3^32) = \frac{\kappa_{11}}{n^3} + \frac{45}{n^2(n-1)}\kappa_9\kappa_2 + \frac{9(17n-26)}{n^2(n-1)^2}\kappa_8\kappa_3$$

$$+ \frac{27(11n^2-31n+22)}{n^2(n-1)^3}\kappa_7\kappa_4 + \frac{9(49n^2-134n+103)}{n^2(n-1)^3}\kappa_6\kappa_5$$

$$+ \frac{54(12n-23)}{n(n-1)^2(n-2)}\kappa_7\kappa_2^2 + \frac{54(63n^2-220n+178)}{n(n-1)^3(n-2)}\kappa_6\kappa_3\kappa_2$$

$$+ \frac{54(93n^2-340n+316)}{n(n-1)^3(n-2)}\kappa_5\kappa_4\kappa_2 + \frac{54(71n^3-421n^2+842n-564)}{n(n-1)^3(n-2)^2}\kappa_5\kappa_3^2$$

$$+ \frac{108(41n^3-257n^2+543n-390)}{n(n-1)^3(n-2)^2}\kappa_4^2\kappa_3 + \frac{108(33n^2-122n+110)}{(n-1)^3(n-2)^2}\kappa_5\kappa_2^3$$

$$+ \frac{648(29n^2-126n+137)}{(n-1)^3(n-2)^2}\kappa_4\kappa_3\kappa_2^2 + \frac{324(29n^2-136n+164)}{(n-1)^3(n-2)^2}\kappa_3^3\kappa_2$$

$$+ \frac{1296n(5n-12)}{(n-1)^3(n-2)^2}\kappa_3\kappa_2^4. \quad (18)$$

METHOD OF CONSTRUCTION OF THE TABLES

8. That part of the tables which express the product of power-sums in terms of the augmented symmetrics (the part below the main diagonal of the tables) was constructed by building up for a given weight from the lower weights. For instance, the expression of (3)(2)(1) in Table 1.6 may be derived in three ways: by multiplying (3)(2) by (1); by multiplying (3)(1) by (2); and by multiplying (2)(1) by (3). From Table 1.3 we have

$$(2)(1) = [3] + [21].$$

Hence
$$(3)(2)(1) = (3)[3]+(3)[21]$$
$$= [6]+[3^2]+[51]+[42]+[321]. \qquad (19)$$

Again, from Table 1.4, $(3)(1) = [4]+[31]$.

Hence
$$(2)(3)(1) = (2)[4]+(2)[31]$$
$$= [6]+[42]+[51]+[3^2]+[321],$$

agreeing with (19).

9. Each line was checked in this manner by being calculated in two different ways by one of us (M.G.K.) and was independently checked by the other (F.N.D.) by symbolic operation. Any product of the one-part functions may be expressed as a sum of monomial symmetrics; for example,

$$(3)(2)(1) = A(6)+B(51)+C(42)+E(3^2)+F(41^2)+G(321)+H(2^3)$$
$$+I(31^3)+J(2^21^2)+K(21^4)+L(1^6), \qquad (20)$$

where A, B, etc., are positive constants to be determined and we do not use D because we require it for a different purpose. In fact, using MacMahon's D-operator technique we have, for example,

$$D_6(3)(2)(1) = 1 = A,$$

all other terms vanishing. Similar operations with D_5D_1, D_4D_2, D_3^2, $D_3D_2D_1$, etc., give

$$A = B = C = G = 1, \quad E = 2,$$

and all other terms zero. Conversion into the [ ] functions then gives equation (19).

This was the method of check actually used. It may be noted that any table of given weight can be built up more speedily from those of lower weight by a similar process, e.g if we operate on equation (20) by D_1 we get

$$(3)(2) = B(5)+F(41)+H(2^3)+J(21^2)+K(1^4),$$

and the values for the coefficients may be read off from the table for weight 5. Operation by D_2 gives

$$(3)(1) = C(4)+G(31)+H(2^2)+J(21^2)+K(1^4),$$

and the values of those coefficients not given by weight 5 are found from the table for weight 4. This method was not used as a check because it only repeats the process described in § 8 and was therefore considered to be insufficient.

10. The complementary part of the tables, expressing the augmented symmetrics in terms of power-sums, was constructed by inverting the relations below the main diagonal. Consider, for instance, Table 1.8. Starting in the cell with row (7)(1) and column [71] we complete the column [71]. We then proceed to the cell in row (6)(2), column [62], and working upwards, complete the column [62]; and so on from column to column. A check is provided by the fact that the number in the top row (8) must be $(-1)^{\rho-1}(\rho-1)!$, where ρ is the number of parts in the item at the head of the column. There are various other relations between the numbers which act as a check on any doubtful figure.

11. Again, an independent check was applied. For example, analogously to equation (20) we have

$$(321) = A'(6) + B'(5)(1) + C'(4)(2) + E'(3)^2 + F'(4)(1)^2 + G'(3)(2)(1) + H'(2)^3 + I'(3)(1)^3 + J'(2)^2(1)^2 + K'(2)(1)^4 + L'(1)^6. \quad (21)$$

By carrying out a series of operations, each time getting zero on the left-hand side, we obtain a set of equations which may be solved for the constant coefficients. This may sound formidable, but in actual practice so many of the coefficients are zero and so many short cuts may be devised that the procedure was not found unmanageable.

12. The present tables are complete for weights of 12 or lower. Complete tables for higher weights would involve a lot of additional labour and more printing space than their usefulness would justify. It may be noted, however, that expressions for higher weights can be obtained from those for lower weights, when necessary, by the symbolic method exemplified in § 9. A similar method can be used for the converse relations.

13. We should like to congratulate the Press on the way in which the tables are arranged and on the uniform excellence of their type-setting.

REFERENCES

FISHER, R. A. (1929). Moments and product-moments of sampling distributions. *Proc. Lond. Math. Soc.* (2), **30**, 199.

KENDALL, M. G. (1941). *The Advanced Theory of Statistics*, **1**. 4th edition, 1948. Charles Griffin and Co.

MACMAHON, P. A. (1915). *Combinatory Analysis*, **1**. Cambridge University Press.

Table 1.2

$w=2$	[2]	$[1^2]$
(2)	1	−1
$(1)^2$	1	1

Table 1.3

$w=3$	[3]	[21]	$[1^3]$
(3)	1	−1	2
(2)(1)	1	1	−3
$(1)^3$	1	3	1

Table 1.4

$w=4$	[4]	[31]	$[2^2]$	$[21^2]$	$[1^4]$
(4)	1	−1	−1	2	−6
(3)(1)	1	1	.	−2	8
$(2)^2$	1	.	1	−1	3
$(2)(1)^2$	1	2	1	1	−6
$(1)^4$	1	4	3	6	1

Table 1.5

$w=5$	[5]	[41]	[32]	$[31^2]$	$[2^21]$	$[21^3]$	$[1^5]$
(5)	1	−1	−1	2	2	−6	24
(4)(1)	1	1	.	−2	−1	6	−30
(3)(2)	1	.	1	−1	−2	5	−20
$(3)(1)^2$	1	2	1	1	.	−3	20
$(2)^2(1)$	1	1	1	.	1	−3	15
$(2)(1)^3$	1	3	4	3	3	1	−10
$(1)^5$	1	5	10	10	15	10	1

Table 1.6

$w=6$	[6]	[51]	[42]	$[41^2]$	$[3^2]$	[321]	$[31^3]$	$[2^3]$	$[2^21^2]$	$[21^4]$	$[1^6]$
(6)	1	−1	−1	2	−1	2	−6	2	−6	24	−120
(5)(1)	1	1	.	−2	.	−1	6	.	4	−24	144
(4)(2)	1	.	1	−1	.	−1	3	−3	5	−18	90
$(4)(1)^2$	1	2	1	1	.	.	−3	.	−1	12	−90
$(3)^2$	1	.	.	.	1	−1	2	.	2	−8	40
(3)(2)(1)	1	1	1	.	1	1	−3	.	−4	20	−120
$(3)(1)^3$	1	3	3	3	1	3	1	.	.	−4	40
$(2)^3$	1	.	3	.	.	.	.	1	−1	3	−15
$(2)^2(1)^2$	1	2	3	3	2	4	.	1	1	−6	45
$(2)(1)^4$	1	4	7	6	4	16	4	3	6	1	−15
$(1)^6$	1	6	15	15	10	60	20	15	45	15	1

Table 1.7

$w=7$	$[7]$	$[61]$	$[52]$	$[51^2]$	$[43]$	$[421]$	$[41^3]$	$[3^21]$	$[32^2]$	$[321^3]$	$[31^4]$	$[2^31]$	$[2^21^3]$	$[21^5]$	$[1^7]$
(7)	1	−1	−1	2	−1	2	−6	3	2	−6	24	−6	24	−120	720
(5) (1)	1	1	.	−2	.	−1	6	−1	.	4	−24	2	−18	120	−840
(5) (2)	1	.	1	−1	.	−1	3	.	−2	3	−12	6	−18	84	−504
$(5)(1)^2$	1	2	1	1	.	.	−3	.	.	−1	12	.	6	−60	504
(4) (3)	1	.	.	.	1	−1	2	−2	−1	4	−14	3	−14	70	−420
(4) (2) (1)	1	1	1	.	1	1	−3	.	.	−2	12	−3	15	−90	630
$(4)(1)^3$	1	3	3	3	1	3	1	.	.	.	−4	.	−1	20	−210
$(3)^2(1)$	1	1	.	.	2	.	.	1	.	−2	8	.	6	−40	280
$(3)(2)^2$	1	.	2	.	1	.	.	.	1	−1	3	−3	8	−35	210
$(3)(2)(1)^2$	1	2	2	1	3	2	.	2	1	1	−6	.	−6	50	−420
$(3)(1)^4$	1	4	6	6	5	12	4	4	3	6	1	.	.	−5	70
$(2)^3(1)$	1	1	3	.	3	3	.	.	3	.	.	1	−3	15	−105
$(2)^2(1)^3$	1	3	5	3	7	9	1	6	7	6	.	3	1	−10	105
$(2)(1)^5$	1	5	11	10	15	35	10	20	25	40	5	15	10	1	−21
$(1)^7$	1	7	21	21	35	105	35	70	105	210	35	105	105	21	1

Table 1.8

$w=8$ (i)	$[8]$	$[71]$	$[62]$	$[61^2]$	$[53]$	$[521]$	$[51^3]$	$[4^2]$	$[431]$	$[42^2]$	$[421^2]$	$[41^4]$
(8)	1	−1	−1	2	−1	2	−6	−1	2	2	−6	24
(7) (1)	1	1	.	−2	.	−1	6	.	−1	.	4	−24
(6) (2)	1	.	1	−1	.	−1	3	.	.	−2	3	−12
$(6)(1)^2$	1	2	1	1	.	.	−3	.	.	.	−1	12
(5) (3)	1	.	.	.	1	−1	2	.	−1	.	2	−8
(5) (2) (1)	1	1	1	.	1	1	−3	.	.	.	−2	12
$(5)(1)^3$	1	3	3	3	1	3	1	.	.	.	.	−4
$(4)^2$	1	.	.	.	.	.	.	1	−1	−1	2	−6
(4) (3) (1)	1	1	.	.	1	.	.	1	1	.	−2	8
$(4)(2)^2$	1	.	2	.	.	.	.	1	.	1	−1	3
$(4)(2)(1)^2$	1	2	2	1	2	2	.	1	2	1	1	−6
$(4)(1)^4$	1	4	6	6	4	12	4	1	4	3	6	1
$(3)^2(2)$	1	.	1	.	2	.	.	.	.	.	.	.
$(3)^2(1)^2$	1	2	1	1	2	.	.	2	4	.	.	.
$(3)(2)^2(1)$	1	1	2	.	3	2	.	1	1	1	.	.
$(3)(2)(1)^3$	1	3	4	3	5	6	1	3	9	3	3	.
$(3)(1)^5$	1	5	10	10	11	30	10	5	25	15	30	5
$(2)^4$	1	.	4	.	6	.	.	3	.	6	.	.
$(2)^3(1)^2$	1	2	4	1	6	6	.	3	6	6	3	.
$(2)^2(1)^4$	1	4	8	6	12	20	4	7	28	16	18	1
$(2)(1)^6$	1	6	16	15	26	66	20	15	90	60	105	15
$(1)^8$	1	8	28	28	56	168	56	35	280	210	420	70

$w=8$ (ii)	$[3^22]$	$[3^21^2]$	$[32^21]$	$[321^3]$	$[31^5]$	$[2^4]$	$[2^31^2]$	$[2^21^4]$	$[21^6]$	$[1^8]$
(8)	2	−6	−6	24	−120	−6	24	−120	720	−5040
(7) (1)	.	4	2	−18	120	.	−12	96	−720	5760
(6) (2)	−1	1	4	−12	60	8	−20	84	−480	3360
$(6)(1)^2$	.	−1	.	6	−60	.	2	−36	360	−3360
(5) (3)	−2	4	4	−14	64	.	−12	64	−384	2688
(5) (2) (1)	.	.	−2	9	−60	.	12	−72	504	−4032
$(5)(1)^3$	.	.	.	−1	20	.	.	8	−120	1344
$(4)^2$	.	2	1	−6	30	3	−6	30	−180	1260
(4) (3) (1)	.	−4	−1	12	−70	.	6	−56	420	−3360
$(4)(2)^2$	.	.	−1	3	−15	−6	9	−33	180	−1260
$(4)(2)(1)^2$	.	.	.	−3	30	.	−3	30	−270	2520
$(4)(1)^4$	.	.	.	.	−5	.	.	−1	30	−420
$(3)^2(2)$	1	−1	−2	5	−20	.	6	−28	100	−1120
$(3)^2(1)^2$	1	1	.	−3	20	.	.	12	−120	1120
$(3)(2)^2(1)$	2	.	1	−3	15	.	−6	32	−210	1680
$(3)(2)(1)^3$	4	3	3	1	−10	.	.	−8	100	−1120
$(3)(1)^5$	10	10	15	10	1	.	.	.	−6	112
$(2)^4$	.	.	.	.	.	1	.	3	−15	105
$(2)^3(1)^2$	6	.	6	.	.	1	1	−6	45	−420
$(2)^2(1)^4$	20	12	28	8	.	3	6	1	−15	210
$(2)(1)^6$	70	60	150	80	6	15	45	15	1	−28
$(1)^8$	280	280	840	560	56	105	420	210	28	1

Table 1.9

w=9 (i)	[9]	[81]	[72]	[71²]	[63]	[621]	[61³]	[54]	[531]	[52²]	[521²]	[51⁴]	[4²1]	[432]	[431²]	[42²1]	[421³]
(9)	1	−1	−1	2	−1	2	−6	−1	2	2	−6	24	2	2	−6	−6	24
(8) (1)	1	1	.	−2	.	−1	6	.	−1	.	4	−24	−1	.	4	2	−18
(7) (2)	1	.	1	−1	.	−1	3	.	.	−2	3	−12	.	−1	1	4	−12
(7) (1)²	1	2	1	1	.	.	−3	.	.	.	−1	12	.	.	−1	.	6
(6) (3)	1	.	.	.	1	−1	2	.	−1	.	2	−8	.	−1	2	2	−8
(6) (2) (1)	1	1	1	.	1	1	−3	.	.	.	−2	12	.	.	.	−2	9
(6) (1)³	1	3	3	3	1	3	1	.	.	.	.	−4	.	.	.	.	−1
(5) (4)	1	.	.	.	.	.	.	1	−1	−1	2	−6	−2	−1	4	3	−12
(5) (3) (1)	1	1	.	.	1	.	.	1	1	.	−2	8	.	.	−2	.	6
(5) (2)²	1	.	2	.	.	.	.	1	.	1	−1	3	.	.	.	−1	3
(5) (2) (1)³	1	2	2	1	2	2	.	1	2	1	1	−6	.	.	.	.	−3
(5) (1)⁴	1	4	6	6	4	12	4	1	4	3	6	1	.	.	.	.	.
(4)³ (1)	1	1	.	.	.	.	.	2	.	.	.	.	1	.	−2	−1	6
(4) (3) (2)	1	.	1	.	1	.	.	1	.	.	.	.	.	1	−1	−2	5
(4) (3) (1)²	1	2	1	1	1	.	.	3	2	.	.	.	2	1	1	.	−3
(4) (2)² (1)	1	1	2	.	2	2	.	2	.	1	.	.	1	2	.	1	−3
(4) (2) (1)⁴	1	3	4	3	4	6	1	4	6	.	3	3	.	3	4	3	3	1
(4) (1)⁵	1	5	10	10	10	30	10	6	20	15	30	5	5	10	10	15	10
(3)³	1	.	.	.	3	.	.	.	.	.	.	.	.	.	.	.	.
(3)³ (2) (1)	1	1	1	.	3	1	.	2	2	.	.	.	.	2	.	.	.
(3)² (1)³	1	3	3	3	3	3	1	6	6	.	.	.	6	6	6	.	.
(3) (2)⁴	1	.	3	.	1	.	.	3	.	3	.	.	.	3	.	.	.
(3) (2)³ (1)²	1	2	3	1	5	4	.	5	6	3	2	.	2	7	1	2	.
(3) (2) (1)⁴	1	4	7	6	9	16	4	11	20	9	12	1	12	23	18	12	4
(3) (1)⁶	1	6	15	15	21	60	20	21	66	45	90	15	30	75	75	90	60
(2)⁴ (1)	1	1	4	.	4	4	.	6	.	6	.	.	3	12	.	6	.
(2)³ (1)³	1	3	6	3	10	12	1	12	18	12	9	.	9	30	9	18	3
(2)² (1)⁵	1	5	12	10	20	40	10	26	60	36	50	5	35	100	70	80	30
(2) (1)⁷	1	7	22	21	42	112	35	56	182	126	231	35	105	350	315	420	245
(1)⁹	1	9	36	36	84	252	84	126	504	378	756	126	315	1260	1260	1890	1260

w=9 (ii)	[41⁵]	[3³]	[3²21]	[3¹2³]	[32³]	[32¹2²]	'[321⁴]	[31⁶]	[2⁴1]	[2³1²]	[2²1⁵]	[21⁷]	[1⁹]
(9)	−120	2	−6	24	−6	24	−120	720	24	−120	720	−5040	40320
(8) (1)	120	.	2	−18	.	−12	96	−720	−6	72	−600	5040	−45360
(7) (2)	60	.	2	−6	6	−14	60	−360	−24	90	−480	3240	−25920
(7) (1)²	−60	.	.	6	.	2	−36	360	.	−18	240	−2520	25920
(6) (3)	40	−3	3	−14	2	−14	64	−360	−8	58	−360	2520	−20160
(6) (2) (1)	−60	.	−1	3	.	8	−48	360	8	−60	420	−3360	30240
(6) (1)³	20	.	1	−1	.	.	8	−120	.	2	−60	840	−10080
(5) (4)	54	.	2	−12	3	−10	54	−324	−12	54	−324	2268	−18144
(5) (3) (1)	−40	.	−2	12	.	8	−56	384	.	−36	320	−2688	24192
(5) (2)²	−15	.	.	.	−3	4	−15	90	12	−36	174	−1134	9072
(5) (2) (1)³	30	.	.	.	.	−2	18	−180	.	18	−180	1764	−18144
(5) (1)⁵	−5	.	.	.	.	.	−1	30	.	.	10	−210	3024
(4)³ (1)	−30	.	.	6	.	2	−24	180	3	−18	150	−1260	11340
(4) (3) (2)	−20	.	−2	6	−3	9	−38	210	12	−51	280	−1890	15120
(4) (3) (1)²	20	.	.	−6	.	−1	24	−210	.	9	−140	1470	−15120
(4) (2)² (1)	15	.	.	.	.	−2	12	−90	−6	27	−105	1260	−11340
(4) (2) (1)³	−10	.	.	.	.	.	−4	60	.	−3	50	−630	7560
(4) (1)⁶	1	.	.	.	.	.	.	−6	.	.	−1	42	−756
(3)³	.	1	−1	2	.	2	−8	40	.	−6	40	−280	2240
(3)² (2) (1)	.	1	1	−3	.	−4	20	−120	.	18	−140	1120	−10080
(3)² (1)⁴	.	1	3	1	.	.	−4	40	.	.	20	−280	3360
(3) (2)³	.	.	.	.	1	−1	3	−15	−4	11	−50	315	−2520
(3) (2)² (1)²	.	2	4	1	1	1	−6	45	.	−9	80	−735	7560
(3) (2) (1)⁴	.	4	16	4	3	6	1	−15	.	.	−10	175	−2520
(3) (1)⁶	6	10	60	20	15	45	15	1	.	.	.	−7	168
(2)⁴ (1)	.	.	.	.	4	.	.	.	1	.	3	15	−105
(2)³ (1)³	.	6	18	.	10	9	.	.	.	3	1	−10	105
(2)² (1)⁵	1	20	100	20	40	70	10	.	15	10	1	−21	378
(2) (1)⁷	21	70	490	140	210	525	140	7	105	105	21	1	−36
(1)⁹	126	280	2520	840	1260	3780	1260	84	945	1260	378	36	1

Table 1.10

$w=10$ (i)	$[10]$	$[91]$	$[82]$	$[81^2]$	$[73]$	$[721]$	$[71^3]$	$[64]$	$[631]$	$[62^2]$	$[621^2]$	$[61^4]$
(10)	1	−1	−1	2	−1	2	−6	−1	2	2	−6	24
(9) (1)	1	1	.	−2	.	−1	6	.	−1	.	4	−24
(8) (2)	1	.	1	−1	.	−1	3	.	.	−2	3	−12
(8) $(1)^2$	1	2	1	1	.	.	−3	.	.	.	−1	12
(7) (3)	1	.	.	.	1	−1	2	.	−1	.	2	−8
(7) (2) (1)	1	1	1	.	1	1	−3	.	.	.	−2	12
(7) $(1)^3$	1	3	3	3	1	3	1	.	.	.	.	−4
(6) (4)	1	.	.	.	.	.	.	1	−1	−1	2	−6
(6) (3) (1)	1	1	.	.	1	.	.	1	1	.	−2	8
(6) $(2)^2$	1	.	2	.	.	.	.	1	.	1	−1	3
(6) (2) $(1)^2$	1	2	2	1	2	2	.	1	2	1	1	−6
(6) $(1)^4$	1	4	6	6	4	12	4	1	4	3	6	1
(5)^2	1	.	.	.	.	.	.	.	.	.	.	.
(5) (4) (1)	1	1	.	.	.	.	.	.	.	.	.	.
(5) (3) (2)	1	.	1	.	1	.	.	1	.	.	.	.
(5) (3) $(1)^2$	1	2	1	1	1	.	.	.	.	.	.	.
(5) $(2)^2$ (1)	1	1	2	.	2	2	.	2	2	.	.	.
(5) (2) $(1)^3$	1	3	3	3	2	2	.	1	.	1	.	.
(5) $(1)^5$	1	5	10	10	10	30	10	3	6	3	3	.
(4)^2 (2)	1	.	1	.	.	.	.	5	20	15	30	5
$(4)^2$ $(1)^2$	1	2	1	1	.	.	.	2	.	.	.	.
$(4)^2$ $(3)^2$	1	.	.	.	2	.	.	2	.	.	.	.
(4) (3) (2) (1)	1	1	1	.	2	1	.	1	.	.	.	.
(4) (3) $(1)^2$	1	3	3	3	2	3	1	2	1	.	.	.
(4) $(2)^3$	1	.	3	.	.	.	.	4	3	.	.	.
(4) $(2)^2$ $(1)^2$	1	2	3	2	4	4	.	4	.	3	.	.
(4) (2) $(1)^4$	1	4	7	6	8	10	4	4	4	3	2	.
(4) $(1)^6$	1	6	15	15	20	60	20	8	16	9	12	1
$(3)^3$ (1)	1	1	.	.	3	.	.	16	60	45	90	15
$(3)^2$ $(2)^2$	1	.	2	.	2	.	.	3	3	.	.	.
$(3)^2$ (2) $(1)^2$	1	2	2	1	4	2	.	1	.	1	.	.
$(3)^2$ $(1)^4$	1	4	6	6	6	12	4	5	6	1	1	.
(3) $(2)^3$ (1)	1	1	3	.	4	3	.	9	12	3	6	1
(3) $(2)^2$ $(1)^3$	1	3	5	3	8	9	1	4	1	3	.	.
(3) (2) $(1)^5$	1	5	11	10	16	35	10	10	15	7	6	.
(3) $(1)^7$	1	7	21	21	36	105	35	20	45	25	40	5
$(2)^5$	1	.	5	.	.	.	.	42	147	105	210	35
$(2)^4$ $(1)^2$	1	2	5	1	8	8	.	10	.	10	.	.
$(2)^3$ $(1)^4$	1	4	9	6	16	24	4	10	8	10	4	.
$(2)^2$ $(1)^6$	1	6	17	15	32	72	20	22	40	24	24	1
(2) $(1)^8$	1	8	29	28	64	176	56	46	120	70	120	15
$(1)^{10}$	1	10	45	45	120	360	120	98	336	238	448	70
								210	840	630	1260	210

$w=10$ (ii)	$[5^2]$	$[541]$	$[532]$	$[531^2]$	$[52^21]$	$[521^3]$	$[51^5]$	$[4^32]$	$[4^31^2]$	$[43^2]$	$[4321]$
(10)	−1	2	2	−6	−6	24	−120	2	−6	2	−6
(9) (1)	.	−1	.	4	2	−18	120	.	4	.	2
(8) (2)	.	.	−1	1	4	−12	60	−1	1	.	2
(8) $(1)^2$	.	.	.	−1	.	6	−60	.	−1	.	.
(7) (3)	.	.	−1	2	−2	−8	40	.	.	−2	3
(7) (2) (1)	.	.	.	.	−2	9	−60	.	.	.	−1
(7) $(1)^3$	.	.	.	.	.	−1	20	.	.	.	.
(6) (4)	.	−1	.	2	1	−6	30	−2	4	−1	3
(6) (3) (1)	.	.	.	−2	.	6	−40	.	.	.	−1
(6) $(2)^2$	.	.	.	.	−1	3	−15	.	.	.	.
(6) (2) $(1)^2$	.	.	.	.	.	−3	30	.	.	.	.
(6) $(1)^4$	.	.	.	.	.	1	−5	.	.	.	.
$(5)^2$	1	−1	−1	2	−2	−6	24	.	2	.	.
(5) (4) (1)	1	1	.	−2	−1	6	−30	.	−4	.	1
(5) (3) (2)	1	.	1	−1	−2	5	−20	.	.	.	−1
(5) (3) $(1)^2$	1	2	1	1	.	−3	20	.	.	.	−1
(5) $(2)^2$ (1)	1	1	2	.	1	−3	15	.	.	.	.
(5) (2) $(1)^3$	1	3	4	3	3	1	−10	.	.	.	.
(5) $(1)^5$	1	5	10	10	15	10	1	.	.	.	.
(4)^2 (2)	.	.	.	.	.	.	.	1	1	.	−1
$(4)^2$ $(1)^2$	2	4	.	.	.	.	.	1	1	.	.
$(4)^2$ $(3)^2$	.	.	.	.	.	.	.	1	.	1	−1
(4) (3) (2) (1)	1	1	1	.	.	.	.	1	.	1	1
(4) (3) $(1)^3$	3	0	3	3	.	.	.	3	3	1	3
(4) $(2)^3$	.	.	.	.	.	.	.	3	.	.	.
(4) $(2)^2$ $(1)^2$	2	4	4	.	2	.	.	3	1	2	4
(4) (2) $(1)^4$	4	10	10	12	12	4	.	7	6	4	10
(4) $(1)^6$	6	36	60	60	90	60	6	15	15	10	60
$(3)^3$ (1)	.	.	.	.	.	.	.	.	.	3	.
$(3)^2$ $(2)^2$	2	.	4	.	.	.	.	.	.	1	.
$(3)^2$ (2) $(1)^2$	2	4	4	2	.	.	.	2	.	5	4
$(3)^2$ $(1)^4$	6	24	12	12	.	.	.	12	12	9	24
(3) $(2)^3$ (1)	3	3	9	.	3	.	.	3	.	3	3
(3) $(2)^2$ $(1)^3$	5	15	19	9	9	2	.	9	3	13	21
(3) (2) $(1)^5$	11	35	61	50	45	20	1	35	30	35	115
(3) $(1)^7$	21	147	231	231	315	210	21	105	105	105	525
$(2)^5$	.	.	.	.	.	.	.	15	.	.	.
$(2)^4$ $(1)^2$	6	12	24	.	12	.	.	15	3	12	24
$(2)^3$ $(1)^4$	12	48	72	36	48	12	.	30	18	48	120
$(2)^2$ $(1)^6$	26	156	232	180	216	100	6	135	105	100	600
(2) $(1)^8$	56	448	784	728	1008	616	56	455	420	360	2800
$(1)^{10}$	126	1260	2520	2520	3780	2520	252	1575	1575	2100	12600

Table 1.10 (*cont.*)

$w=10$ (iii)	$[431^3]$	$[42^3]$	$[42^31^2]$	$[421^4]$	$[41^5]$	$[3^31]$	$[3^32^2]$	$[3^321^2]$	$[3^31^4]$	$[32^21]$
(10)	24	−6	24	−120	720	−6	−6	24	−120	24
(9) (1)	−18	.	−12	96	−720	2	.	−12	96	−6
(8) (2)	−6	6	−14	60	−360	.	4	−8	36	−18
(8) (1)^2	6	.	2	−36	360	.	.	2	−36	.
(7) (3)	−8	.	−8	40	−240	6	4	−16	64	−12
(7) (2) (1)	3	.	8	−48	360	.	.	4	−24	6
(7) (1)^3	−1	.	.	8	−120	.	.	.	8	.
(6) (4)	−12	5	−12	54	−300	3	1	−10	54	−8
(6) (3) (1)	6	.	4	−32	240	−3	.	10	−56	2
(6) (2)^2	.	−3	4	−15	90	.	−1	1	−3	6
(6) (2) (1)^2	.	.	−2	18	−180	.	.	−1	6	.
(6) (1)^4	.	.	.	−1	30	.	.	.	−1	.
(5)^2	−6	.	−4	24	−144	.	2	−4	24	−6
(5) (4) (1)	12	.	6	−48	324	.	.	4	−48	3
(5) (3) (2)	3	.	4	−20	120	.	−4	6	−24	12
(5) (3) (1)^2	−3	.	.	12	−120	.	.	−2	24	.
(5) (2)^2 (1)	.	.	−2	12	−90	.	.	.	.	−3
(5) (2) (1)^3	.	.	.	−4	60	.	.	.	.	.
(5) (1)^5	.	.	.	.	−6	.	.	.	.	.
(4)^2 (2)	3	−3	5	−18	90	.	.	2	−12	3
(4)^2 (1)^2	−3	.	−1	12	−90	.	.	.	12	.
(4) (3)^2	2	.	2	−8	40	−3	−1	6	−22	3
(4) (3) (2) (1)	−3	.	−4	20	−120	.	.	−4	24	−3
(4) (3) (1)^3	1	.	.	−4	40	.	.	.	−8	.
(4) (2)^3	.	1	−1	3	−15	.	.	.	.	1
(4) (2)^2 (1)^2	.	1	1	−6	45	.	.	.	.	.
(4) (2) (1)^4	4	3	6	1	−15	.	.	.	.	.
(4) (1)^6	20	15	45	15	1	.	.	.	.	.
(3)^4 (1)	.	.	.	.	.	1	.	−2	8	.
(3)^3 (2)^2	.	.	.	.	.	.	1	−1	3	−3
(3)^3 (2) (1)^3	.	.	.	.	.	2	1	1	−6	.
(3)^3 (1)^6	8	.	.	.	.	4	3	6	1	.
(3) (2)^5 (1)	.	1	.	.	.	3	3	.	.	1
(3) (2)^3 (1)^3	1	3	3	.	.	6	7	6	.	3
(3) (2) (1)^5	30	15	30	5	.	20	25	40	5	15
(3) (1)^9	175	105	315	105	7	70	105	210	35	105
(2)^5	.	10	.	.	.	.	.	.	.	.
(2)^4 (1)^2	.	10	6	.	.	.	12	.	.	8
(2)^3 (1)^6	12	28	36	3	.	24	48	36	.	40
(2)^2 (1)^8	140	120	240	45	1	120	220	300	30	240
(2) (1)^8	840	630	1680	490	28	560	1120	1960	280	1680
(1)^{10}	4200	3150	9450	3150	210	2800	6300	12600	2100	12600

$w=10$ (iv)	$[32^21^3]$	$[321^5]$	$[31^7]$	$[2^5]$	$[2^41^2]$	$[2^31^4]$	$[2^21^6]$	$[21^8]$	$[1^{10}]$
(10)	−120	720	−5040	24	−120	720	−5040	40320	−362880
(9) (1)	72	−600	5040	.	48	−480	4320	−40320	403200
(8) (2)	66	−360	2520	−30	102	−504	3240	−25200	226800
(8) (1)^2	−18	240	−2520	.	−6	144	−1800	20160	−226800
(7) (3)	64	−360	2400	.	48	−336	2400	−19200	172800
(7) (2) (1)	−42	300	−2520	.	−48	360	−2880	25920	−259200
(7) (1)^3	2	−60	840	.	.	−24	480	−6720	80400
(6) (4)	48	−300	2100	−20	52	−300	2100	−16800	151200
(6) (3) (1)	−42	320	−2520	.	−16	232	−2160	20160	−201600
(6) (2)^2	−18	90	−630	20	−44	186	−1110	8400	−75600
(6) (2) (1)^3	12	−120	1260	.	8	−120	1260	−13440	151200
(6) (1)^6	.	10	−210	.	.	2	−90	1680	−25200
(5)^3	24	−144	1008	.	24	−144	1008	−8064	72576
(5) (4) (1)	−30	270	−2268	.	−24	216	−1944	18144	−181440
(5) (3) (2)	−40	204	−1344	.	−48	264	−1728	13440	−120960
(5) (3) (1)^3	12	−140	1344	.	.	−72	960	−10752	120960
(5) (2)^2 (1)	12	−75	630	.	24	−144	1044	−9072	90720
(5) (2) (1)^4	−2	30	−420	.	.	24	−360	4704	−60480
(5) (1)^5	.	−1	42	.	.	.	12	−336	6048
(4)^2 (2)	−15	90	−630	15	−27	126	−810	6300	−56700
(4)^2 (1)^3	3	−60	630	.	3	−36	450	−5040	56700
(4) (3)^2	−20	110	−700	.	−12	96	−700	5600	−50400
(4) (3) (2) (1)	27	−190	1470	.	24	−204	1680	−15120	151200
(4) (3) (1)^3	−1	40	−490	.	.	12	−280	3920	−50400
(4) (2)^3	3	−15	105	−10	14	−51	285	−2100	18900
(4) (2)^2 (1)^2	−3	30	−315	.	−6	54	−495	5040	−56700
(4) (2) (1)^4	.	−5	105	.	.	−3	75	−1260	18900
(4) (1)^6	.	.	−7	.	.	.	−1	56	−1260
(3)^3 (1)	6	−40	280	.	.	−24	240	−2240	22400
(3)^2 (2)^2	8	−35	210	.	12	−60	370	−2800	25200
(3)^2 (2) (1)^3	−6	50	−420	.	.	36	−420	4480	−50400
(3)^2 (1)^6	.	−5	70	.	.	.	30	−560	8400
(3) (2)^4 (1)	−3	15	−105	.	−8	44	−300	2520	−25200
(3) (2)^3 (1)^3	1	−10	105	.	.	−12	160	−1960	25200
(3) (2) (1)^5	10	1	−21	.	.	.	−12	280	−5040
(3) (1)^7	105	21	1	.	.	.	.	−8	240
(2)^5	.	.	.	1	−1	3	−15	105	−945
(2)^4 (1)^2	.	.	.	1	1	−6	45	−420	4725
(2)^3 (1)^4	12	.	.	3	6	1	−15	210	−3150
(2)^2 (1)^6	140	12	.	15	45	15	1	−28	630
(2) (1)^8	1400	224	8	105	420	210	28	1	−45
(1)^{10}	12600	2520	120	945	4725	3150	630	45	1

Table 1.11

$w-11$ (i)	[11]	[10, 1]	[92]	[91²]	[83]	[821]	[81³]	[74]	[731]	[72²]	[721²]	[71⁴]	[65]	[641]	[632]
(11)	1	−1	−1	2	−1	2	−6	−1	2	2	−6	24	−1	2	2
(10) (1)	1	1	.	−2	.	−1	6	.	−1	.	4	−24	.	−1	.
(9) (2)	1	.	1	−1	.	−1	3	.	.	−2	3	−12	.	.	−1
(9) (1)²	1	2	1	1	.	.	−3	.	.	.	−1	12	.	.	.
(8) (3)	1	.	.	.	1	−1	2	.	−1	.	2	−8	.	.	−1
(8) (2) (1)	1	1	1	.	1	1	−3	.	.	.	−2	12	.	.	.
(8) (1)³	1	3	3	3	1	3	1	.	.	.	.	−4	.	.	.
(7) (4)	1	.	.	.	.	.	.	1	−1	−1	2	−6	.	−1	.
(7) (3) (1)	1	1	.	.	1	.	.	1	1	.	−2	8	.	.	.
(7) (2)²	1	.	2	.	.	.	.	1	.	1	−1	3	.	.	.
(7) (2) (1)³	1	2	2	1	2	2	.	1	2	1	1	−6	.	.	.
(7) (1)⁴	1	4	6	6	4	12	4	1	4	3	6	1	.	.	.
(6) (5)	1	.	.	.	.	.	.	1	.	.	.	.	1	−1	−1
(6) (4) (1)	1	1	.	.	.	.	.	1	.	.	.	.	1	1	.
(6) (3) (2)	1	.	1	.	1	.	.	.	.	.	.	.	1	.	1
(6) (3) (1)²	1	2	1	1	1	.	.	2	2	.	.	.	1	2	1
(6) (2)² (1)	1	1	2	.	2	2	.	1	.	1	.	.	1	1	2
(6) (2) (1)³	1	3	4	3	4	6	1	3	6	3	3	.	1	3	4
(6) (1)⁴	1	5	10	10	10	30	10	5	20	15	30	.	5	5	10
(5)² (1)	1	1	.	.	.	.	.	.	.	.	.	.	2	.	.
(5) (4) (2)	1	.	1	.	.	.	.	1	.	.	.	.	1	.	.
(5) (4) (1)²	1	2	1	1	.	.	.	1	.	.	.	.	3	2	.
(5) (3)²	1	.	.	.	2	.	.	.	.	.	.	.	1	.	.
(5) (3) (2) (1)	1	1	1	.	2	1	.	1	1	.	.	.	2	.	1
(5) (3) (1)³	1	3	3	3	2	3	1	3	3	.	.	.	4	6	3
(5) (2)³	1	.	3	.	.	.	.	3	.	3	.	.	1	.	.
(5) (2)² (1)²	1	2	3	1	4	4	.	3	4	3	2	.	3	2	4
(5) (2) (1)⁴	1	4	7	6	8	16	4	7	16	9	12	1	5	12	16
(5) (1)⁵	1	6	15	15	20	60	20	15	60	45	90	15	7	30	60
(4)² (3)	1	.	.	.	1	.	.	2	.	.	.	.	.	.	.
(4)² (2) (1)	1	1	1	.	1	1	.	2	.	.	.	.	2	2	.
(4)² (1)⁴	1	3	3	3	1	3	1	2	.	.	.	.	6	6	.
(4) (3)² (1)	1	1	.	.	2	.	.	3	2	.	.	.	1	1	.
(4) (3) (2)²	1	.	2	.	1	.	.	2	.	1	1	.	.	.	2
(4) (3) (2) (1)³	1	2	2	1	3	2	.	4	4	1	1	.	2	4	2
(4) (3) (1)⁴	1	4	6	6	5	12	4	6	8	3	6	1	10	16	6
(4) (2)³ (1)	1	1	3	.	3	3	.	4	.	3	.	.	4	4	6
(4) (2)² (1)³	1	3	5	3	7	9	1	8	12	7	6	.	8	12	14
(4) (2) (1)⁵	1	5	11	10	15	35	10	16	40	25	40	5	16	40	50
(4) (1)⁷	1	7	21	21	35	105	35	36	140	105	210	35	28	112	210
(3)³ (2)	1	.	1	.	3	.	.	.	.	.	.	.	3	.	3
(3)³ (1)²	1	2	1	1	3	.	.	6	6	.	.	.	3	6	3
(3)² (2)² (1)	1	1	2	.	4	2	.	3	2	1	.	.	5	1	6
(3)² (2) (1)³	1	3	4	3	6	6	1	9	12	3	3	.	9	15	12
(3)² (1)⁶	1	5	10	10	12	30	10	15	30	15	30	5	21	45	30
(3) (2)⁴	1	.	4	.	1	.	.	6	.	6	.	.	4	.	4
(3) (2)³ (1)²	1	2	4	1	7	6	.	8	8	6	3	.	10	8	16
(3) (2)² (1)⁴	1	4	8	6	13	20	4	18	32	16	18	1	20	40	48
(3) (2) (1)⁶	1	6	16	15	27	66	20	36	96	60	105	15	42	120	136
(3) (1)⁸	1	8	28	28	57	168	56	78	288	210	420	70	84	336	588
(2)⁵ (1)	1	1	5	.	5	5	.	10	.	10	.	.	10	10	20
(2)⁴ (1)²	1	3	7	3	13	15	1	18	24	18	12	.	22	30	52
(2)³ (1)⁴	1	5	13	10	25	45	10	38	80	48	60	5	46	110	160
(2)² (1)⁶	1	7	23	21	49	119	35	78	224	148	252	35	98	322	504
(2) (1)⁸	1	9	37	36	93	261	84	162	576	414	792	126	210	882	1596
(1)¹¹	1	11	55	55	165	495	165	330	1320	990	1980	330	462	2310	4620

Table 1.11 (*cont.*)

w = 11 (ii)	[631⁴]	[62²1]	[621³]	[61⁵]	[5²1]	[542]	[541³]	[53²]	[5321]	[531³]	[52³]	[52²1³]	[521⁴]	[51⁶]	[4²3]
(11)	−6	−6	24	−120	2	2	−6	2	−6	24	−6	24	−120	720	2
(10) (1)	4	2	−18	120	−1	.	4	.	2	−18	.	−12	96	−720	.
(9) (2)	1	4	−12	60	.	−1	1	.	2	−6	6	−14	60	−360	.
(9) (1)²	−1	.	6	−60	.	.	−1	.	.	6	.	2	−36	360	.
(8) (3)	2	2	−8	40	.	.	.	−2	3	−8	.	−8	40	−240	−1
(8) (2) (1)	.	−2	9	−60	.	.	.	.	−1	3	.	8	−48	360	.
(8) (1)³	.	.	−1	20	.	.	.	.	.	−1	.	.	8	−120	.
(7) (4)	2	1	−6	30	.	−1	2	.	1	−6	3	−6	30	−180	−2
(7) (3) (1)	−2	.	6	−40	.	.	.	.	−1	6	.	4	−32	240	.
(7) (2)²	.	−1	3	−15	.	.	.	.	.	.	−3	4	−15	90	.
(7) (2) (1)³	.	.	−3	30	.	.	.	.	.	.	.	−2	18	−180	.
(7) (1)⁴	.	.	.	−5	.	.	.	.	.	.	.	.	−1	30	.
(6) (5)	2	2	−6	24	−2	−1	4	−1	3	−12	2	−10	48	−264	.
(6) (4) (1)	−2	−1	6	−30	.	.	−2	.	.	6	.	2	−24	180	.
(6) (3) (2)	−1	−2	5	−20	.	.	.	.	−1	3	.	4	−20	120	.
(6) (3) (1)³	1	.	−3	20	.	.	.	.	.	−3	.	.	12	−120	.
(6) (2)² (1)	.	1	−3	15	.	.	.	.	.	.	.	−2	12	−90	.
(6) (2) (1)³	3	3	1	−10	.	.	.	.	.	.	.	.	−4	60	.
(6) (1)⁴	10	15	10	1	.	.	.	.	.	.	.	4	−24	144	.
(5)³ (1)	.	.	.	.	1	.	−2	.	−1	6	.	.	.	−6	.
(5) (4) (2)	.	.	.	.	.	1	−1	.	−1	3	−3	5	−18	90	.
(5) (4) (1)²	.	.	.	.	2	1	1	.	.	−3	.	−1	12	−90	.
(5) (3)²	.	.	.	.	.	.	.	1	−1	2	.	2	−8	40	.
(5) (3) (2) (1)	.	.	.	.	1	1	.	1	1	−3	.	−4	20	−120	.
(5) (3) (1)³	3	.	.	.	3	3	3	1	3	1	.	.	−4	40	.
(5) (2)³	.	.	.	.	.	3	.	.	.	.	1	−1	3	−15	.
(5) (2)² (1)²	4	2	.	.	2	3	1	2	4	.	1	1	−6	45	.
(5) (2) (1)³	10	.	.	.	.	4	3	1	6	1	.	3	1	−15	.
(5) (1)⁵	60	90	60	6	6	15	15	10	60	20	15	45	15	1	.
(4)³ (3)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
(4)³ (2) (1)	.	.	.	.	.	2	.	.	.	.	.	.	.	.	.
(4)³ (1)³	.	.	.	.	6	6	6	.	.	.	.	.	.	.	1
(4) (3)³ (1)	.	.	.	.	.	.	.	1	.	.	.	.	.	.	1
(4) (3) (2)²	.	.	.	.	.	2	.	.	.	.	.	.	.	.	2
(4) (3) (2) (1)³	1	.	.	.	2	4	1	2	2	.	.	.	.	.	1
(4) (3) (1)⁴	6	.	.	.	12	18	18	4	12	4	.	.	.	.	3
(4) (2)³ (1)²	.	3	.	.	.	6	.	.	.	.	1	.	.	.	3
(4) (2)² (1)⁴	6	9	2	.	6	14	6	6	12	.	3	3	.	.	3
(4) (2) (1)⁵	40	45	20	1	20	46	40	20	80	20	15	30	5	.	7
(4) (1)⁶	210	315	210	21	42	126	126	70	420	140	105	315	105	7	35
(3)³ (2)	.	.	.	.	.	.	.	3	.	.	.	.	.	.	.
(3)³ (1)³	3	.	.	.	.	.	.	3	.	.	.	.	.	.	6
(3)² (2)² (1)	.	1	.	.	.	4	.	5	4	.	.	.	.	.	2
(3)² (2) (1)³	9	3	1	.	6	12	6	9	12	2	.	.	.	.	12
(3)² (1)⁵	30	15	10	1	30	60	60	21	60	20	.	.	.	.	30
(3) (2)⁴	.	.	.	.	.	12	.	.	.	.	.	.	.	.	3
(3) (2)³ (1)²	1	6	.	.	6	18	3	12	18	.	4	.	.	.	9
(3) (2)² (1)⁴	30	28	8	.	20	52	30	32	76	12	12	18	2	.	35
(3) (2) (1)⁶	135	150	80	6	66	186	165	96	366	100	60	135	30	1	105
(3) (1)⁷	588	840	560	56	168	588	588	336	1848	616	420	1260	420	28	315
(2)⁵ (1)	.	10	.	.	.	30	.	.	.	.	10	.	.	.	15
(2)⁴ (1)³	12	30	4	.	18	66	18	36	72	.	22	18	.	.	39
(2)³ (1)⁵	100	120	40	1	60	198	120	120	360	60	76	120	15	.	135
(2)² (1)⁷	420	532	280	21	182	658	546	392	1624	420	336	756	175	7	455
(2) (1)⁸	1512	2142	1344	126	504	2142	2016	1344	7056	2184	1638	4536	1386	84	1575
(1)¹¹	4620	6930	4620	462	1386	6930	6930	4620	27720	9240	6930	20790	6930	462	5775

Table 1.11 (*cont.*)

w = 11 (iii)	$[4^2 21]$	$[4^3 1^3]$	$[43^2 1]$	$[432^2]$	$[4321^3]$	$[431^4]$	$[42^2 1]$	$[42^2 1^3]$	$[421^4]$	$[41^7]$	$[3^4 2]$	$[3^3 1^3]$	$[3^3 2^2 1]$	$[3^3 21^3]$	$[3^2 1^5]$
(11)	−6	24	−6	−6	24	−120	24	−120	720	−5040	−6	24	24	−120	720
(10) (1)	2	−18	2	.	−12	96	−6	72	−600	5040	.	−12	−6	72	−600
(9) (2)	2	−6	.	.	2	−36	−18	66	−360	2520	2	−2	−12	42	−240
(9) (1)^2	.	6	.	4	−8	36	.	−18	240	−2520	.	2	.	−18	240
(8) (3)	1	−2	4	2	−10	40	−6	40	−240	1680	6	−18	−16	70	−360
(8) (2) (1)	−1	3	.	.	4	−24	6	−42	300	−2520	.	.	4	−24	180
(8) (1)^3	.	−1	.	.	.	8	.	2	−60	840	.	.	.	2	−60
(7) (4)	4	−12	4	3	−12	54	−12	54	−300	1980	.	−12	−6	48	−300
(7) (3) (1)	.	.	−2	.	6	−32	6	−24	200	−1680	.	12	4	−48	320
(7) (2)^2	.	.	.	−1	1	−3	6	−18	90	−630	.	.	2	−6	30
(7) (2) (1)^3	.	.	.	.	−1	6	.	12	−120	1260	.	.	.	6	−60
(7) (1)^4	.	.	.	.	.	−1	.	.	10	−210	.	.	.	.	10
(6) (5)	2	−12	1	2	−8	48	−8	42	−264	1848	3	−6	−10	42	−264
(6) (4) (1)	−2	12	−1	.	6	−48	5	−36	270	−2100	.	6	1	−30	270
(6) (3) (2)	.	.	.	−2	3	−12	6	−22	120	−840	−3	3	10	−20	140
(6) (3) (1)^2	.	.	.	.	−1	12	.	6	−80	840	.	−3	.	15	−140
(6) (2)^2 (1)	.	.	.	.	.	.	−3	12	−75	630	.	.	−1	3	−15
(6) (2) (1)^3	.	.	.	.	.	.	.	−2	30	−420	.	.	.	−1	10
(6) (1)^5	.	.	.	.	.	.	.	.	−1	42	.	.	.	.	−1
(5)^2 (1)	.	6	.	.	2	−24	.	−12	120	−1008	.	.	2	−12	120
(5) (4) (2)	−2	6	.	−2	5	−24	9	−33	174	−1134	.	.	4	−18	120
(5) (4) (1)^2	.	−6	.	.	−1	24	.	9	−120	1134	.	.	.	6	−120
(5) (3)^2	.	.	−1	.	2	−8	.	−6	40	−280	−3	6	6	−22	104
(5) (3) (2) (1)	.	.	.	.	−2	12	.	12	−100	840	.	.	−4	18	−120
(5) (3) (1)^3	.	.	.	.	.	−4	.	.	20	−280	.	.	.	−2	40
(5) (2)^3	.	.	.	.	.	.	−1	3	−15	105	.	.	.	.	.
(5) (2)^2 (1)^2	.	.	.	.	.	.	.	−3	30	−315	.	.	.	.	.
(5) (2) (1)^4	.	.	.	.	.	.	.	.	−5	105	.	.	.	.	.
(5) (1)^6	.	.	.	.	.	.	.	.	.	−7	.	.	.	.	.
(4)^2 (3)	−1	2	−2	−1	4	−14	3	−14	70	−420	.	6	2	−18	100
(4)^2 (2) (1)	1	−3	.	.	−2	12	−3	15	−90	630	.	.	.	6	−60
(4)^2 (1)^3	3	1	.	.	.	−4	.	−1	20	−210	.	.	.	.	20
(4) (3)^2 (1)	.	.	1	.	−2	8	.	6	−40	280	.	−6	−1	18	−110
(4) (3) (2)^2	.	.	.	1	−1	3	−3	8	−35	210	.	.	−2	6	−30
(4) (3) (2) (1)^2	2	.	2	1	1	−6	.	−6	50	−420	.	.	.	−6	60
(4) (3) (1)^4	12	4	4	3	6	1	.	.	−5	70	.	.	.	.	−10
(4) (2)^3 (1)	3	.	.	3	.	.	1	−3	15	−105	.	.	.	.	.
(4) (2)^2 (1)^3	9	1	6	7	6	.	3	1	−10	105	.	.	.	.	.
(4) (2) (1)^5	35	10	20	25	40	5	15	10	1	−21	.	.	.	.	.
(4) (1)^7	105	35	70	105	210	35	105	105	21	1	.	.	.	.	.
(3)^3 (2)	.	.	.	.	.	.	.	.	.	.	1	−1	−2	5	−20
(3)^3 (1)^2	.	.	6	.	.	.	.	.	.	.	1	1	.	−3	20
(3)^2 (2)^2 (1)	.	.	1	2	.	.	.	.	.	.	2	.	1	−3	15
(3)^2 (2) (1)^3	6	.	15	6	6	.	.	.	.	.	4	3	3	1	−10
(3)^2 (1)^4	60	20	45	30	60	10	.	.	.	.	10	10	15	10	1
(3) (2)^4	.	.	.	6	.	.	.	.	.	.	.	.	.	.	.
(3) (2)^3 (1)^2	6	.	6	12	3	.	2	.	.	.	6	.	6	.	.
(3) (2)^2 (1)^4	36	4	52	44	42	1	12	4	.	.	.	.	.	8	.
(3) (2) (1)^6	210	60	210	210	345	45	90	60	6	.	20	12	28	.	.
(3) (1)^9	840	280	840	1050	2100	350	840	840	168	8	70	60	150	80	6
(2)^5 (1)	15	.	.	30	.	.	10	.	.	.	280	280	840	560	56
(2)^4 (1)^3	45	3	36	78	36	.	30	6	.	.	.	.	.	.	.
(2)^3 (1)^5	195	30	240	300	300	15	140	60	3	.	24	.	36	.	.
(2)^2 (1)^7	945	245	1120	1400	2100	245	840	560	63	1	120	60	240	60	.
(2) (1)^9	4095	1260	5040	6930	12600	1890	5670	5040	882	36	560	420	1540	700	42
(1)^{11}	17325	5775	23100	34650	69300	11550	34650	34650	6930	330	2800	2520	10080	5880	504
											15400	15400	69300	46200	4620

Table 1.11 (*cont.*)

$w=11$ (iv)	$[32^4]$	$[32^31^2]$	$[32^21^4]$	$[321^7]$	$[31^9]$	$[2^51]$	$[2^41^3]$	$[2^31^5]$	$[2^21^7]$	$[21^9]$	$[1^{11}]$
(11)	24	−120	720	−5040	40320	−120	720	−5040	40320	−362880	3628800
(10) (1)	.	48	−480	4320	−40320	24	−360	3600	−35280	362880	−3991680
( ) (2)	−24	78	−384	2520	−20160	120	−552	3360	−25200	221760	−2217600
(9) (1)^2	.	−6	144	−1800	20160	.	72	−1200	15120	−181440	2217600
(8) (3)	−6	60	−360	2400	−18480	30	−300	2280	−18480	166320	−1663200
(8) (2) (1)	.	−36	264	−2160	20160	−30	306	−2520	22680	−226800	2494800
(8) (1)^3	.	.	−24	480	−6720	.	−6	240	−4200	60480	−831600
(7) (4)	−12	42	−276	1980	−15840	60	−288	1980	−15840	142560	−1425600
(7) (3) (1)	.	−24	256	−2160	19200	.	144	−1680	16800	−172800	1900800
(7) (2)^2	12	−24	102	−630	5040	−60	216	−1170	8280	−71280	712800
(7) (2) (1)^3	.	6	−84	900	−10080	.	−72	900	−10080	116640	−1425600
(7) (1)^4	.	.	2	−90	1680	.	.	−30	840	−15120	237600
(6) (5)	−8	46	−264	1848	−14784	40	−264	1848	−14784	133056	−1330560
(6) (4) (1)	.	−16	192	−1800	18000	−20	150	−1500	14700	−151200	1663200
(6) (3) (2)	8	−44	212	−1320	10080	−40	256	−1660	12600	−110880	1108800
(6) (3) (1)^3	.	2	−84	960	−10080	.	−24	580	−7560	90720	−1108800
(6) (2)^3 (1)	.	12	−72	540	−5040	20	−132	930	−7770	75600	−831600
(6) (2) (1)^4	.	.	16	−240	3360	.	8	−200	2940	−40320	554400
(6) (1)^6	.	.	.	12	−336	.	.	2	−126	3024	−55440
(5)^3 (1)	.	−12	96	−864	8064	.	72	−720	7056	−72576	798336
(5) (4) (2)	12	−33	168	−1134	9072	−60	252	−1512	11340	−99792	997920
(5) (4) (1)^2	.	3	−60	810	−9072	.	−36	540	−6804	81648	−997920
(5) (3)^2	.	−18	104	−664	4928	.	72	−600	4928	−44352	443520
(5) (3) (2) (1)	.	24	−160	1224	−10752	.	−144	1320	−12096	120960	−1330560
(5) (3) (1)^3	.	.	16	−280	3584	.	.	−120	2240	−32256	443520
(5) (2)^3	−4	5	−18	105	−840	20	−60	294	−1974	16632	−166320
(5) (2)^2 (1)^3	.	−3	24	−225	2520	.	36	−360	3654	−40824	498960
(5) (2) (1)^5	.	.	−2	45	−840	.	.	30	−630	10584	−166320
(5) (1)^8	.	.	.	−1	56	.	.	.	14	−504	11088
(4)^2 (3)	3	−12	86	−600	4620	−15	78	−570	4620	−41580	415800
(4)^2 (2) (1)	.	6	−60	540	−5040	15	−81	630	−5670	56700	−623700
(4)^2 (1)^3	.	.	4	−120	1680	.	3	−60	1050	−15120	207900
(4) (3)^2 (1)	.	6	−80	660	−5600	.	−36	480	−4900	50400	−554400
(4) (3) (2)^2	−6	15	−65	390	−2940	30	−120	675	−4830	41580	−415800
(4) (3) (2) (1)^3	.	−3	54	−570	5880	.	36	−510	5880	−68040	831600
(4) (3) (1)^4	.	.	−1	60	−980	.	.	15	−490	8820	−138600
(4) (2)^3 (1)	.	−2	12	−90	840	−10	42	−255	1995	−18900	207900
(4) (2)^2 (1)^3	.	.	−4	60	−840	.	−6	90	−1155	15120	−207900
(4) (2) (1)^5	.	.	.	−6	168	.	.	−3	105	−2268	41580
(4) (1)^7	.	.	.	.	−8	.	.	.	−1	72	−1980
(3)^3 (2)	.	6	−28	160	−1120	.	−24	180	−1400	12320	−123200
(3)^3 (1)^2	.	.	12	−120	1120	.	.	−60	840	−10080	123200
(3)^2 (2)^2 (1)	.	−6	32	−210	1680	.	36	−300	2590	−25200	277200
(3)^2 (2) (1)^3	.	.	−8	100	−1120	.	.	60	−980	13440	−184800
(3)^2 (1)^6	.	.	.	−6	112	.	.	.	42	−1008	18480
(3) (2)^4	1	−1	3	−15	105	−5	14	−65	420	−3465	34650
(3) (2)^3 (1)^3	1	1	−6	45	−420	.	−12	110	−1050	11340	−138600
(3) (2)^2 (1)^5	3	6	1	−15	210	.	.	−15	280	−4410	69300
(3) (2) (1)^7	15	45	15	1	−28	.	.	.	−14	420	−9240
(3) (1)^9	105	420	210	28	1	.	.	.	.	−9	330
(2)^5 (1)	5	.	.	.	.	1	−3	15	−105	945	−10395
(2)^4 (1)^3	13	12	.	.	.	3	1	−10	105	−1260	17325
(2)^3 (1)^5	55	100	15	.	.	15	10	1	−21	378	−6930
(2)^2 (1)^7	315	840	245	14	.	105	105	21	1	−36	990
(2) (1)^9	2205	7560	3150	336	9	945	1260	378	36	1	−55
(1)^{11}	17325	69300	34650	4620	165	10395	17325	6930	990	55	1

Table 1.12

$w=12$ (i)	[12]	[11, 1]	[10, 2]	[10, 1^2]	[93]	[921]	[91^3]	[84]	[831]	[82^2]	[821^2]	[81^4]	[75]	[741]
(12)	1	−1	−1	2	−1	2	−6	−1	2	2	−6	24	−1	2
(11) (1)	1	1	.	−2	.	−1	6	.	−1	.	4	−24	.	−1
(10) (2)	1	.	1	−1	.	−1	3	.	.	−2	3	−12	.	.
(10) $(1)^2$	1	2	1	1	.	.	−3	.	.	.	−1	12	.	.
(9) (3)	1	.	.	.	1	−1	2	.	−1	.	2	−8	.	.
(9) (2) (1)	1	1	1	.	1	1	−3	.	.	.	−2	12	.	.
(9) $(1)^3$	1	3	3	3	1	3	1	.	.	.	.	−4	.	.
(8) (4)	1	.	.	.	.	.	.	1	−1	−1	2	−6	.	−1
(8) (3) (1)	1	1	.	.	1	.	.	1	1	.	−2	8	.	.
(8) $(2)^2$	1	.	2	.	.	.	.	1	.	1	−1	3	.	.
(8) (2) $(1)^2$	1	2	2	1	2	2	.	1	2	1	1	−6	.	.
(8) $(1)^4$	1	4	6	6	4	12	4	1	4	3	6	1	.	.
(7) (5)	1	.	.	.	.	.	.	1	.	.	.	.	1	−1
(7) (4) (1)	1	1	.	.	.	.	.	1	.	.	.	.	1	1
(7) (3) (2)	1	.	1	.	1	.	.	.	.	.	.	.	1	.
(7) (3) $(1)^2$	1	2	1	1	1	.	.	2	2	.	.	.	1	2
(7) $(2)^2$ (1)	1	1	2	.	2	2	.	1	.	1	.	.	1	1
(7) (2) $(1)^3$	1	3	4	3	4	6	1	3	6	3	3	.	1	3
(7) $(1)^5$	1	5	10	10	10	30	10	5	20	15	30	5	1	5
$(6)^2$	1	.	.	.	.	.	.	.	.	.	.	.	.	.
(6) (5) (1)	1	1	.	.	.	.	.	.	.	.	.	.	1	.
(6) (4) (2)	1	.	1	.	.	.	.	1	.	.	.	.	.	.
(6) (4) $(1)^2$	1	2	1	1	.	.	.	1	.	.	.	.	1	.
(6) $(3)^2$	1	.	.	.	2	.	.	.	.	.	.	.	2	2
(6) (3) (2) (1)	1	1	1	.	2	1	.	1	1	.	.	.	.	.
(6) (3) $(1)^3$	1	3	3	3	2	3	1	3	3	.	.	.	1	.
(6) $(2)^3$	1	1	3	.	3	.	.	3	.	3	.	.	3	6
(6) $(2)^2$ $(1)^2$	1	2	3	1	4	4	.	3	4	3	.	.	2	.
(6) (2) $(1)^4$	1	4	7	6	8	10	4	7	16	9	12	1	4	12
(6) $(1)^6$	1	6	15	15	20	60	20	15	60	45	90	15	6	30
(5) $(2)^2$	1	.	1	.	.	.	.	.	.	.	.	.	2	.
(5) $(2)^2$ $(1)^2$	1	2	1	1	.	.	.	.	.	.	.	.	2	.
(5) (4) (3)	1	.	.	.	1	.	.	1	.	.	.	.	1	.
(5) (4) (2) (1)	1	1	1	.	1	1	.	1	.	.	.	.	1	.
(5) (4) $(1)^3$	1	3	3	3	1	3	1	1	.	.	.	.	2	1
(5) $(3)^2$ (1)	1	1	.	.	2	.	.	2	2	.	.	.	4	3
(5) (3) $(2)^2$	1	.	2	.	1	.	.	1	.	1	.	.	1	.
(5) (3) (2) $(1)^3$	1	2	2	1	3	2	.	3	4	1	1	.	3	.
(5) (3) $(1)^4$	1	4	6	6	5	12	4	5	8	3	6	1	3	2
(5) $(2)^2$ (1)	1	1	3	.	3	3	.	3	.	3	.	.	1	1
(5) $(2)^3$ $(1)^2$	1	3	5	3	7	9	1	7	12	7	6	.	7	12
(5) (2) $(1)^4$	1	5	11	10	15	35	10	15	40	25	40	5	4	3
(5) $(1)^6$	1	7	21	21	35	105	35	35	140	105	210	35	6	9
$(4)^3$	1	.	.	.	.	.	.	3	.	.	.	.	12	35
$(4)^3$ (3) (1)	1	1	.	.	1	.	.	3	1	.	.	.	22	105
$(4)^3$ $(2)^2$	1	.	2	.	.	.	.	3	.	1	.	.	.	.
$(4)^3$ (2) $(1)^2$	1	2	2	1	2	2	.	3	.	1	.	.	2	2
$(4)^3$ $(1)^4$	1	4	6	6	4	12	4	3	2	1	1	.	.	.
(4) $(3)^2$ (2)	1	2	1	1	2	.	.	3	4	3	6	1	4	4
(4) $(3)^2$ $(1)^2$	1	1	2	.	2	.	.	1	.	.	.	.	8	8
(4) (3) $(2)^2$ (1)	1	1	2	.	3	2	.	5	4	.	.	.	2	.
(4) (3) (2) $(1)^3$	1	3	4	3	5	6	1	7	9	3	3	.	4	6
(4) (3) $(1)^5$	1	5	10	10	11	30	10	11	25	15	30	.	8	2
(4) $(2)^4$	1	.	1	.	.	.	.	7	.	6	.	5	16	12
(4) $(2)^3$ $(1)^2$	1	2	4	1	6	6	.	7	6	6	3	.	.	30
(4) $(2)^2$ $(1)^4$	1	4	8	6	12	20	4	15	28	16	18	1	8	8
(4) (2) $(1)^6$	1	6	16	15	26	66	20	31	90	60	105	15	16	32
(4) $(1)^8$	1	8	28	28	56	168	56	71	280	210	420	70	32	96
$(3)^4$	1	.	1	.	4	.	.	.	.	.	.	.	64	288
$(3)^3$ (2) (1)	1	1	1	.	4	1	.	3	3	.	.	.	.	.
$(3)^3$ $(1)^3$	1	3	3	3	4	3	1	9	9	.	.	.	3	.
$(3)^2$ $(2)^2$	1	.	3	.	2	.	.	3	.	3	.	.	9	18
$(3)^2$ $(2)^2$ $(1)^2$	1	2	3	1	6	6	.	7	8	3	3	.	6	.
$(3)^2$ (2) $(1)^4$	1	4	7	6	10	16	4	15	24	9	12	1	8	6
$(3)^2$ $(1)^6$	1	6	15	15	22	60	20	27	72	45	90	15	18	36
(3) $(2)^4$ (1)	1	1	4	.	5	4	.	7	6	6	.	.	36	90
(3) $(2)^3$ $(1)^3$	1	3	6	3	11	12	1	15	21	12	9	.	10	6
(3) $(2)^2$ $(1)^5$	1	5	12	10	21	40	10	31	65	36	50	5	18	24
(3) (2) $(1)^7$	1	7	22	21	43	112	35	63	189	126	231	35	38	90
(3) $(1)^9$	1	9	36	36	85	252	84	135	513	378	756	126	78	252
$(2)^6$	1	.	6	.	.	.	.	15	.	15	.	.	162	702
$(2)^5$ $(1)^2$	1	2	6	1	10	10	.	15	10	15	5	.	.	.
$(2)^4$ $(1)^4$	1	4	10	6	20	28	4	31	52	33	30	1	20	20
$(2)^3$ $(1)^6$	1	6	18	15	38	78	20	63	150	93	135	15	40	72
$(2)^2$ $(1)^8$	1	8	30	28	72	184	56	127	392	267	476	70	84	228
(2) $(1)^{10}$	1	10	46	45	130	370	120	255	930	675	1305	210	176	624
$(1)^{12}$	1	12	66	66	220	660	220	495	1980	1485	2970	495	372	1620
													792	3960

Table 1.12 (*cont.*)

w = 12 (ii)	[732]	[7³1²]	[7²51]	[721⁴]	[71⁵]	[6²]	[651]	[642]	[641⁴]	[63³]	[6321]	[631³]	[62²]	[62²1⁴]
(12)	2	−6	−6	24	−120	−1	2	2	−6	2	−6	24	−6	24
(11)(1)	.	4	2	−18	120	.	−1	.	4	.	2	−18	.	−12
(10)(2)	−1	1	4	−12	60	.	.	−1	1	.	2	−6	6	−14
(10)(1)²	.	−1	.	6	−60	.	.	.	−1	.	.	6	.	2
(9)(3)	−1	2	2	−8	40	.	.	.	.	−2	3	−8	.	−8
(9)(2)(1)	.	.	−2	9	−60	.	.	.	.	.	−1	3	.	8
(9)(1)³	.	.	.	−1	20	.	.	.	.	.	.	−1	.	.
(8)(4)	.	2	1	−6	30	.	.	−1	2	.	1	−6	3	−6
(8)(3)(1)	.	−2	.	6	−40	.	.	.	.	.	−1	6	.	4
(8)(2)²	.	.	−1	3	−15	.	.	.	.	.	.	.	−3	4
(8)(2)(1)²	.	.	.	−3	30	.	.	.	.	.	.	.	.	−2
(8)(1)⁴	.	.	.	.	−5	.	.	.	.	.	.	.	.	.
(7)(5)	−1	2	2	−6	24	.	−1	.	2	.	1	−6	.	−4
(7)(4)(1)	.	−2	−1	6	−30	.	.	.	−2	.	.	6	.	2
(7)(3)(2)	1	−1	−2	5	−20	.	.	.	.	.	−1	3	.	4
(7)(3)(1)³	1	1	.	−3	20	.	.	.	.	.	.	−3	.	.
(7)(2)²(1)	1	.	1	−3	15	.	.	.	.	.	.	.	.	−2
(7)(2)(1)³	4	3	3	1	−10	.	.	.	.	.	.	.	.	.
(7)(1)⁵	10	10	15	10	1	.	.	.	.	.	.	.	.	.
(6)³	.	.	.	.	.	1	−1	−1	2	−1	2	−6	2	−6
(6)(5)(1)	.	.	.	.	.	1	1	.	−2	.	−1	6	.	4
(6)(4)(2)	.	.	.	.	.	1	.	1	−1	.	−1	3	−3	5
(6)(4)(1)²	.	.	.	.	.	1	2	1	1	.	.	−3	.	−1
(6)(3)²	.	.	.	.	.	1	.	.	.	1	−1	3	.	2
(6)(3)(2)(1)	1	.	.	.	.	1	1	1	.	1	1	−3	.	−4
(6)(3)(1)³	3	3	.	.	.	1	3	3	3	1	3	1	.	.
(6)(2)³	.	.	.	.	.	1	.	3	.	.	.	.	1	−1
(6)(2)²(1)²	4	.	2	.	.	1	2	3	1	2	4	.	1	1
(6)(2)(1)⁴	16	12	12	4	.	1	4	7	6	4	16	4	3	6
(6)(1)⁶	60	60	90	60	6	1	6	15	15	10	60	20	15	45
(5)²(2)	.	.	.	.	.	.	.	.	.	.	.	.	.	.
(5)²(1)²	.	.	.	.	.	2	4	.	.	.	.	.	.	.
(5)(4)(3)	.	.	.	.	.	.	.	.	.	.	.	.	.	.
(5)(4)(2)(1)	.	.	.	.	.	1	1	1	.	.	.	.	.	.
(5)(4)(1)³	.	.	.	.	.	3	9	3	3	.	.	.	.	.
(5)(3)²(1)	.	.	.	.	.	1	1	.	.	1	.	.	.	.
(5)(3)(2)²	2	.	.	.	.	.	.	.	.	.	.	.	.	.
(5)(3)(2)(1)²	2	1	.	.	.	2	4	2	.	2	2	.	.	.
(5)(3)(1)³	6	6	.	.	.	4	16	12	12	4	12	4	.	.
(5)(2)³(1)	6	.	3	.	.	1	1	3	.	.	.	.	1	.
(5)(2)²(1)³	14	6	9	2	.	3	9	9	3	6	12	.	3	3
(5)(2)(1)⁴	50	40	45	20	1	5	25	35	30	20	80	20	15	30
(5)(1)⁷	210	210	315	210	21	7	49	105	105	70	420	140	105	315
(4)²	.	.	.	.	.	.	.	.	.	.	.	.	.	.
(4)²(3)(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.
(4)²(2)²	.	.	.	.	.	2	.	4	.	.	.	.	.	.
(4)²(2)(1)³	.	.	.	.	.	2	4	4	2	.	.	.	.	.
(4)²(1)⁴	.	.	.	.	.	6	24	12	12	.	.	.	.	.
(4)(3)²(2)	2	.	.	.	.	1	.	1	.	1	.	.	.	.
(4)(3)²(1)²	2	2	.	.	.	1	2	1	1	1	.	.	.	.
(4)(3)(2)²(1)	4	.	1	.	.	2	2	4	.	2	2	.	.	.
(4)(3)(2)(1)³	8	6	3	1	.	4	12	10	6	4	6	1	.	.
(4)(3)(1)⁴	20	20	15	10	1	10	50	40	40	10	30	10	.	.
(4)(2)³	.	.	.	.	.	4	.	16	.	.	.	.	4	.
(4)(2)²(1)²	12	6	6	.	.	4	8	16	4	6	12	.	4	2
(4)(2)(1)⁴	40	24	28	8	.	8	32	40	24	20	56	8	12	18
(4)(2)(1)⁶	140	120	150	80	6	16	96	136	120	70	300	80	60	135
(4)(1)⁸	560	560	840	560	56	28	224	448	448	280	1680	560	420	1260
(3)³	.	.	.	.	.	3	.	.	.	6	.	.	.	.
(3)²(2)(1)	3	.	.	.	.	3	3	3	.	6	3	.	.	.
(3)²(1)³	9	9	.	.	.	3	9	9	9	6	9	3	.	.
(3)²(2)²	6	.	.	.	.	1	.	3	.	1	.	.	1	.
(3)²(2)²(1)²	10	2	2	.	.	5	10	11	1	11	12	.	1	1
(3)²(2)(1)⁴	30	24	12	4	.	9	36	39	30	21	48	12	3	6
(3)²(1)⁶	90	90	90	60	6	21	126	135	135	51	180	60	15	45
(3)(2)³(1)	16	.	6	.	.	4	4	16	.	4	4	.	4	.
(3)(2)²(1)³	36	12	18	3	.	10	30	42	12	28	48	1	10	9
(3)(2)²(1)⁴	112	80	80	30	1	20	100	140	100	80	240	50	40	70
(3)(2)(1)⁷	372	336	420	245	21	42	294	462	420	252	1092	315	210	525
(3)(1)⁹	1296	1296	1890	1260	126	84	756	1512	1512	924	5292	1764	1260	3780
(2)⁶	.	.	.	.	.	10	.	60	.	.	.	.	20	.
(2)⁴(1)⁴	40	.	20	.	.	10	20	60	10	20	40	.	20	10
(2)⁴(1)⁶	112	48	72	16	.	22	88	148	60	88	208	16	52	60
(2)³(1)⁸	336	240	288	120	6	46	276	468	330	280	960	200	196	360
(2)²(1)¹⁰	1024	896	1184	672	56	98	784	1484	1288	896	4032	1120	868	2128
(2)(1)¹²	3000	2880	4140	2640	252	210	2100	4620	4410	2940	15960	5040	3780	10710
(1)¹²	7920	7920	11880	7920	792	462	5544	13860	13860	9240	55440	18480	13860	41580

Table 1.12 (*cont.*)

$w=12$ (iii)	$[6^2 1^4]$	$[61^6]$	$[5^2 2]$	$[5^2 1^2]$	$[543]$	$[5421]$	$[541^3]$	$[53^2 1]$	$[532^2]$	$[5321^3]$	$[531^4]$	$[52^2 1]$	$[52^2 1^2]$	$[521^5]$
(12)	−120	720	2	−6	2	−6	24	−6	−6	24	−120	24	−120	720
(11) (1)	96	−720	.	4	.	2	−18	2	.	−12	96	−6	72	−600
(10) (2)	60	−360	−1	1	.	2	−6	.	4	−8	36	−18	66	−360
(10) $(1)^2$	−36	360	.	−1	.	.	6	.	.	2	−36	.	−18	240
(9) (3)	40	−240	.	.	−1	1	−2	4	2	−10	40	−6	40	−240
(9) (2) (1)	−48	360	.	.	.	−1	3	.	.	4	−24	6	−42	300
(9) $(1)^3$	8	−120	.	.	.	.	−1	.	.	.	8	.	2	−60
(8) (4)	30	−180	.	.	−1	2	−6	2	1	−6	30	−6	30	−180
(8) (3) (1)	−32	240	.	.	.	.	.	−2	.	6	−32	.	−24	200
(8) $(2)^2$	−15	90	.	.	.	.	.	.	−1	1	−3	6	−18	90
(8) (2) $(1)^2$	18	−180	.	.	.	.	.	.	.	−1	6	.	12	−120
(8) $(1)^4$	−1	30	.	.	.	.	.	.	.	.	−1	.	.	10
(7) (5)	24	−144	−2	4	−1	3	−12	2	4	−10	48	−12	48	−264
(7) (4) (1)	−24	180	.	.	.	−1	6	.	.	2	−24	3	−18	150
(7) (3) (2)	−20	120	.	.	.	.	.	.	−2	3	−12	6	−22	120
(7) (3) $(1)^2$	12	−120	.	.	.	.	.	.	.	−1	12	.	6	−80
(7) $(2)^2$ (1)	12	−90	.	.	.	.	.	.	.	.	.	−3	12	−75
(7) (2) $(1)^3$	−4	60	.	.	.	.	.	.	.	.	.	.	−2	30
(7) $(1)^5$	.	−6	.	.	.	.	.	.	.	.	.	.	.	−1
$(6)^2$	24	−120	.	2	.	1	−6	1	.	−4	24	−2	18	−120
(6) (5) (1)	−24	144	.	−4	.	−1	12	−1	.	6	−48	2	−30	240
(6) (4) (2)	−18	90	.	.	.	−1	3	.	.	2	−12	3	−15	90
(6) (4) $(1)^2$	12	−90	.	.	.	.	−3	.	.	.	12	.	3	−60
(6) $(3)^2$	−8	40	.	.	.	.	.	−1	.	2	−8	.	−6	40
(6) (3) (2) (1)	20	−120	.	.	.	.	.	.	.	−2	12	.	12	−100
(6) (3) $(1)^3$	−4	40	.	.	.	.	.	.	.	.	−4	.	.	20
(6) $(2)^3$	3	−15	.	.	.	.	.	.	.	.	.	−1	3	−15
(6) $(2)^2$ $(1)^2$	−6	45	.	.	.	.	.	.	.	.	.	.	−3	30
(6) (2) $(1)^4$	1	−15	.	.	.	.	.	.	.	.	.	.	.	−5
(6) $(1)^6$	15	1	.	.	.	.	.	.	.	.	.	.	.	1
(5)^2 (2)	.	.	1	−1	.	−1	3	.	−2	3	−12	6	−18	84
(5)^2 $(1)^2$	.	.	1	1	.	.	−3	.	.	−1	12	.	6	−60
(5) (4) (3)	.	.	.	.	1	−1	2	−2	−1	4	−14	3	−14	70
(5) (4) (2) (1)	.	.	1	.	1	1	−3	.	.	−2	12	−3	15	−90
(5) (4) $(1)^3$	.	.	3	3	1	3	1	.	.	.	−4	.	−1	20
(5) (3)^2 (1)	.	.	.	.	2	.	.	1	.	−2	8	.	6	−40
(5) (3) $(2)^2$	.	.	2	1	1	1	.	.	1	−1	3	−3	8	−33
(5) (3) (2) $(1)^3$	.	.	2	1	3	2	.	2	1	1	−6	.	−6	50
(5) (3) $(1)^4$	.	.	6	6	5	12	4	4	3	6	1	.	.	−5
(5) $(2)^3$ (1)	.	.	3	.	3	3	.	.	3	.	.	1	−3	15
(5) $(2)^3$ $(1)^2$	.	.	5	3	7	9	1	6	7	6	.	3	1	−10
(5) (2) $(1)^4$	5	.	11	10	15	35	10	20	25	40	5	15	10	1
(5) $(1)^6$	105	7	21	21	35	105	35	70	105	210	35	105	105	21
$(4)^3$	.	.	.	.	.	.	.	.	.	.	.	.	.	.
$(4)^2$ (3) (1)	.	.	.	.	2	.	.	.	.	.	.	.	.	.
$(4)^2$ $(2)^2$	.	.	.	.	.	.	.	.	.	.	.	.	.	.
$(4)^2$ (2) $(1)^2$	.	.	1	.	4	4	.	.	.	.	.	.	.	.
$(4)^2$ $(1)^4$	.	.	12	12	8	24	8	.	.	.	.	.	.	.
(4) $(3)^2$ (2)	.	.	.	.	2	.	.	.	.	.	.	.	.	.
(4) $(3)^2$ $(1)^2$	.	.	.	.	6	.	.	2	.	.	.	.	.	.
(4) (3) (2) $(1)^2$	.	.	2	.	4	2	.	.	1	.	.	.	.	.
(4) (3) $(2)^2$ (1)	.	.	6	3	14	12	1	6	3	3	.	.	.	.
(4) (3) $(1)^4$	.	.	30	30	30	90	30	20	15	30	5	.	.	.
(4) $(2)^4$	.	.	.	.	.	.	.	.	.	.	.	.	.	.
(4) $(2)^3$ $(1)^2$	.	.	6	.	12	12	.	.	6	.	.	3	.	.
(4) $(2)^2$ $(1)^4$	2	.	20	12	40	56	8	24	28	24	.	12	4	.
(4) (2) $(1)^6$	30	1	66	60	116	276	80	120	150	240	30	90	60	6
(4) $(1)^8$	420	28	168	168	336	1008	336	560	840	1680	280	840	840	168
$(3)^4$	.	.	.	.	.	.	.	.	.	.	.	.	.	.
$(3)^3$ (2) (1)	.	.	.	.	6	.	.	3	.	.	.	.	.	.
$(3)^3$ $(1)^3$	.	.	.	.	18	.	.	9	.	.	.	.	.	.
$(3)^2$ $(2)^3$	.	.	6	.	6	.	.	.	6	.	.	.	.	.
$(3)^2$ $(2)^2$ $(1)^2$	.	.	6	2	18	8	.	10	6	4	.	.	.	.
$(3)^2$ (2) $(1)^4$	1	.	18	12	54	48	8	36	18	24	2	.	.	.
$(3)^2$ $(1)^6$	15	1	90	90	162	360	120	180	90	180	30	.	.	.
(3) $(2)^4$ (1)	.	.	12	.	18	12	.	.	18	.	.	4	.	.
(3) $(2)^3$ $(1)^3$	.	.	24	9	60	54	3	36	42	27	.	12	3	.
(3) $(2)^2$ $(1)^4$	10	.	72	50	186	260	50	160	156	190	15	60	30	2
(3) (2) $(1)^6$	140	7	252	231	588	1302	385	672	756	1281	175	420	315	42
(3) $(1)^9$	1260	84	756	756	1890	5292	1764	3024	4158	8316	1386	3780	3780	756
$(2)^6$	.	.	.	.	.	.	.	.	.	.	.	.	.	.
$(2)^5$ $(1)^2$	.	.	30	.	60	60	.	.	60	.	.	20	.	.
$(2)^4$ $(1)^4$	4	.	84	36	216	264	24	144	216	144	.	88	24	.
$(2)^3$ $(1)^6$	60	1	258	180	708	1188	240	720	888	1080	90	456	240	18
$(2)^2$ $(1)^8$	560	28	840	728	2352	5264	1456	3136	4032	6496	840	2688	2016	280
(2) $(1)^{10}$	3360	210	2646	2520	7980	21420	6720	13440	18900	35280	5460	16380	15120	2772
$(1)^{12}$	13860	924	8316	8316	27720	83160	27720	55440	83160	166320	27720	83160	83160	16632

Table 1.12 (*cont.*)

$w=12$ (iv)	$[5^{12}]$	$[4^9]$	$[4^831]$	$[4^622]$	$[4^521^2]$	$[4^41^4]$	$[43^32]$	$[43^21^3]$	$[432^21]$	$[4321^5]$	$[431^7]$	$[42^6]$	$[42^31^6]$
(12)	−5040	2	−6	−6	24	−120	−6	24	24	−120	720	24	−120
(11) (1)	5040	.	2	.	−12	96	.	−12	−6	72	−600	.	48
(10) (2)	2520	.	.	4	−8	36	2	−2	−12	42	−240	−24	78
$(10)(1)^2$	−2520	.	.	.	2	−36	.	2	.	−18	240	.	−6
(9) (3)	1080	.	2	.	−4	16	4	−12	−10	46	−240	.	36
(9) (2) (1)	−2520	.	.	.	4	−24	.	.	4	−24	180	.	−36
$(9)(1)^3$	840	.	.	.	.	8	.	.	.	2	−60	.	.
(8) (4)	1260	−3	5	5	−14	54	2	−14	−10	54	−300	−18	54
(8) (3) (1)	−1680	.	−1	.	2	−8	.	8	2	−30	200	.	−12
$(8)(2)^2$	−630	.	.	−1	1	−3	.	.	2	−6	30	12	−24
$(8)(2)(1)^2$	1260	.	.	.	−1	6	.	.	.	6	−60	.	6
$(8)(1)^4$	−210	.	.	.	.	−1	.	.	.	.	10	.	.
(7) (5)	1728	.	2	.	−8	48	2	−8	−8	42	−264	.	36
(7) (4) (1)	−1260	.	−2	.	8	−48	.	8	3	−36	270	.	−24
(7) (3) (2)	−840	.	.	.	.	.	−2	2	6	−17	80	.	−24
$(7)(3)(1)^2$	840	.	.	.	.	.	.	−2	.	9	−80	.	.
$(7)(2)^2(1)$	630	.	.	.	.	.	.	.	−1	3	−15	.	12
$(7)(2)(1)^3$	−420	.	.	.	.	.	.	.	.	−1	10	.	.
$(7)(1)^5$	42	.	.	.	.	.	.	.	.	.	−1	.	.
$(6)^2$	840	.	.	2	−4	24	1	−2	−4	18	−120	−8	22
(6) (5) (1)	−1848	.	.	.	4	−48	.	2	2	−24	240	.	−16
(6) (4) (2)	−630	.	.	−4	6	−24	−1	1	6	−21	120	20	−41
$(6)(4)(1)^2$	630	.	.	.	−2	24	.	−1	.	9	−120	.	5
$(6)(3)^2$	−280	.	.	.	.	.	−1	2	2	−8	40	.	−6
(6) (3) (2) (1)	840	.	.	.	.	.	.	.	−2	9	−60	.	12
$(6)(3)(1)^3$	−280	.	.	.	.	.	.	.	.	−1	10	.	.
$(6)(2)^3$	105	.	.	.	.	.	.	.	.	.	.	−4	3
$(6)(2)^2(1)^2$	−315	.	.	.	.	.	.	.	.	.	.	.	−3
$(6)(2)(1)^4$	105	.	.	.	.	.	.	.	.	.	.	.	.
$(6)(1)^6$	−7	.	.	.	.	.	.	.	.	.	.	.	.
$(5)^2(2)$	−504	.	.	.	2	−12	.	.	2	−9	60	.	−12
$(5)^2(1)^2$	504	.	.	.	.	12	.	.	.	3	−60	.	.
(5) (4) (3)	−420	.	−2	.	4	−16	−2	8	5	−26	134	.	−18
(5) (4) (2) (1)	630	.	.	.	−4	24	.	.	−2	15	−120	.	18
$(5)(4)(1)^3$	−210	.	.	.	.	−8	.	.	.	−1	40	.	.
$(5)(3)^2(1)$	280	.	.	.	.	.	.	−2	.	6	−40	.	.
$(5)(3)(2)^2$	210	.	.	.	.	.	.	.	−1	3	−15	.	6
$(5)(3)(2)(1)^3$	−420	.	.	.	.	.	.	.	.	−3	30	.	.
$(5)(3)(1)^5$	70	.	.	.	.	.	.	.	.	.	−5	.	.
$(5)(2)^2(1)$	−105	.	.	.	.	.	.	.	.	.	.	.	−2
$(5)(2)^2(1)^3$	105	.	.	.	.	.	.	.	.	.	.	.	.
$(5)(2)(1)^5$	−21	.	.	.	.	.	.	.	.	.	.	.	.
$(5)(1)^7$	1	.	.	.	.	.	.	.	.	.	.	.	.
$(4)^3$	.	1	−1	−1	2	−6	.	2	1	−6	30	3	−6
$(4)^3(3)(1)$	.	1	1	.	−2	8	.	−4	−1	12	−70	.	6
$(4)^3(2)^2$	.	1	.	1	−1	3	.	.	−1	3	−15	−6	9
$(4)^3(2)(1)^3$	.	1	2	1	1	−6	.	.	.	−3	30	.	−3
$(4)^3(1)^6$	.	1	4	3	6	1	.	.	.	.	−5	.	.
$(4)(3)^3(2)$	.	.	.	.	.	.	1	−1	−2	5	−20	.	6
$(4)(3)^3(1)^3$	.	2	4	.	.	.	1	1	.	−3	20	.	.
$(4)(3)(2)^3(1)$	.	1	1	1	.	.	2	.	1	−3	15	.	−6
$(4)(3)(2)(1)^5$	.	3	9	3	3	.	4	3	3	1	−10	.	.
$(4)(3)(1)^6$	.	5	25	15	30	5	10	10	15	10	1	.	.
$(4)(2)^4$	.	3	.	6	.	.	.	.	.	.	.	1	−1
$(4)(2)^3(1)^2$	.	3	6	6	3	.	6	.	6	.	.	1	1
$(4)(2)^2(1)^4$	.	7	28	16	18	1	20	12	28	8	.	3	6
$(4)(2)(1)^6$	1	15	90	60	105	15	70	60	150	80	6	15	45
$(4)(1)^8$	8	35	280	210	420	70	280	280	840	560	56	105	420
$(3)^4$	.	.	.	.	.	.	.	.	.	.	.	.	.
$(3)^3(2)(1)$	.	.	.	.	.	.	3	.	.	.	.	.	.
$(3)^3(1)^3$	.	6	18	.	.	.	9	9	.	.	.	.	.
$(3)^2(2)^3$	.	.	.	.	.	.	3	.	.	.	.	.	.
$(3)^2(2)^2(1)^3$	.	2	4	2	.	.	11	1	4	.	.	.	.
$(3)^2(2)(1)^5$	.	12	48	12	12	.	39	30	24	8	.	.	.
$(3)^2(1)^6$	.	30	180	90	180	30	135	135	180	120	12	.	.
$(3)(2)^4(1)$	.	3	3	6	.	.	12	.	6	.	.	1	.
$(3)(2)^3(1)^3$	.	9	27	18	9	.	48	9	36	3	.	3	3
$(3)(2)^2(1)^5$	.	35	175	80	90	5	200	130	220	70	1	15	30
$(3)(2)(1)^7$	1	105	735	420	735	105	840	735	1470	805	63	105	315
$(3)(1)^9$	36	315	2835	1890	3780	630	3780	3780	9450	6300	630	945	3780
$(2)^6$	.	15	.	45	.	.	.	.	.	.	.	15	.
$(2)^5(1)^2$	.	15	30	45	15	.	60	.	60	.	.	15	10
$(2)^4(1)^4$	.	39	156	123	90	3	264	72	312	48	.	43	60
$(2)^3(1)^6$	.	135	810	495	585	45	1200	720	1800	600	18	195	420
$(2)^2(1)^8$	8	455	3640	2345	3780	490	5600	4480	11200	5600	392	1155	3360
$(2)(1)^{10}$	120	1575	15750	11025	20475	3150	27300	25200	69300	42000	3780	7875	28350
$(1)^{12}$	792	5775	69300	51975	103950	17325	138600	138600	415800	277200	277200	51975	207900

Table 1.12 (*cont.*)

$w=12$ (v)	$[42^21^4]$	$[421^6]$	$[41^8]$	$[3^4]$	$[3^221]$	$[3^21^6]$	$[3^12^3]$	$[3^12^21^2]$	$[3^121^4]$	$[3^11^8]$	$[32^41]$	$[32^31^2]$	$[32^21^4]$
(12)	720	−5040	40320	−6	24	−120	24	−120	720	−5040	−120	720	−5040
(11) (1)	−480	4320	−40320	.	−6	72	.	48	−480	4320	24	−360	3600
(10) (2)	−384	2520	−20160	.	−6	18	−18	54	−264	1800	96	−432	2640
(10) $(1)^2$	144	−1800	20160	.	.	−18	.	−6	144	−1800	.	72	−1200
(9) (3)	−240	1680	−13440	8	−20	76	−12	72	−384	2400	48	−348	2400
(9) (2) (1)	264	−2160	20160	.	2	−6	.	−24	168	−1440	−24	234	−1920
(9) $(1)^3$	−24	480	−6720	.	.	2	.	.	−24	480	.	−6	240
(8) (4)	−300	1080	−15120	.	−6	54	−6	38	−276	1980	42	−252	1880
(8) (3) (1)	160	−1440	13440	.	6	−54	.	−32	280	−2160	−6	180	−1800
(8) $(2)^2$	102	−630	5040	.	.	.	6	−10	42	−270	−36	126	−690
(8) (2) $(1)^2$	−84	900	−10080	.	.	.	.	4	−48	540	.	−54	660
(8) $(1)^4$	2	−90	1680	.	.	.	.	.	2	−90	.	.	−30
(7) (5)	−240	1728	−13824	.	−6	36	−12	40	−240	1728	48	−252	1728
(7) (4) (1)	216	−1800	15840	.	.	−36	.	−12	192	−1800	−12	126	−1380
(7) (3) (2)	128	−840	6720	.	6	−18	12	−36	160	−960	−48	228	−1360
(7) (3) $(1)^2$	−48	600	−6720	.	.	18	.	4	−96	960	.	−36	640
(7) $(2)^2$ (1)	−72	540	−5040	.	.	.	.	4	−24	180	12	−72	510
(7) (2) $(1)^3$	16	−240	3360	.	.	.	.	.	8	−120	.	6	−140
(7) $(1)^5$	.	12	−336	.	.	.	.	.	.	12	.	.	2
$(6)^2$	−120	840	−6720	3	−6	18	−2	22	−120	840	16	−120	840
(6) (5) (1)	168	−1584	14784	.	3	−18	.	−20	168	−1584	−8	138	−1320
(6) (4) (2)	180	−1110	8400	.	3	−9	3	−21	114	−810	−32	168	−1080
(6) (4) $(1)^2$	−72	810	−8400	.	.	9	.	1	−60	810	.	−24	480
(6) $(3)^2$	40	−280	2240	−6	9	−24	2	−24	112	−640	−8	94	−640
(6) (3) (2) (1)	−88	720	−6720	.	−18	9	.	20	−116	840	8	−132	1060
(6) (3) $(1)^3$	8	−160	2240	.	.	−3	.	.	20	−280	.	2	−140
(6) $(2)^3$	−18	105	−840	.	.	.	−1	1	−3	15	8	−24	120
(6) $(2)^2$ $(1)^2$	24	−225	2520	.	.	.	.	−1	6	−45	.	18	−180
(6) (2) $(1)^4$	−2	45	−840	.	.	.	.	.	−1	15	.	.	30
(6) $(1)^6$	.	−1	56	.	.	.	.	.	.	−1	.	.	.
(5)^2 (2)	72	−504	4032	.	.	.	6	−10	48	−360	−24	90	−528
(5)^2 $(1)^2$	−24	360	−4032	.	.	.	.	2	−24	360	.	−18	240
(5) (4) (3)	120	−804	6048	.	6	−36	6	−28	172	−1128	−24	150	−1064
(5) (4) (2) (1)	−132	1044	−9072	.	.	.	.	8	−72	720	12	−99	840
(5) (4) $(1)^3$	12	−240	3024	.	.	.	.	.	8	−240	.	3	−100
(5) $(3)^2$ (1)	−24	240	−2240	.	−3	18	.	12	−88	624	.	−54	520
(5) (3) $(2)^2$	−32	210	−1680	.	.	.	−6	8	−30	180	24	−78	404
(5) (3) (2) $(1)^2$	24	−300	3360	.	.	.	.	−4	36	−360	.	36	−400
(5) (3) $(1)^4$	.	30	−560	.	.	.	.	.	−2	60	.	.	20
(5) $(2)^3$ (1)	12	−90	840	.	.	.	.	.	.	.	−4	15	−90
(5) $(2)^2$ $(1)^3$	−4	60	−840	.	.	.	.	.	.	.	.	−3	40
(5) (2) $(1)^5$	.	−8	168	.	.	.	.	.	.	.	.	.	−2
(5) $(1)^7$	.	.	−8	.	.	.	.	.	.	.	.	.	.
$(4)^3$	30	−180	1260	.	−6	.	.	−2	24	−180	−3	18	−150
$(4)^2$ (3) (1)	−36	420	−3360	.	18	.	.	4	−72	600	3	−36	430
(4) $(3)^2$ (2)	−33	180	−1260	.	.	.	.	2	−12	90	6	−27	165
$(4)^2$ (3) $(1)^2$	30	−270	2520	.	.	.	.	.	12	−180	.	9	−150
(4) $(3)^2$ $(1)^2$	−1	30	−420	.	.	.	.	.	.	30	.	.	5
(4) (3) $(2)^2$ (2)	−28	160	−1120	.	−3	9	−3	13	−58	330	12	−69	420
(4) (3) $(2)^2$ $(1)^2$	12	−120	1120	.	.	−9	.	−1	36	−330	.	9	−200
(4) (3) (2) $(1)^3$	32	−210	1680	.	.	.	.	−4	24	−180	−6	45	−325
(4) (3) (2) $(1)^3$	−8	100	−1120	.	.	.	.	.	−8	120	.	−3	90
(4) (3) $(1)^4$	.	−6	112	.	.	.	.	.	.	−12	.	.	−1
(4) $(2)^4$	3	−15	105	.	.	.	.	.	.	.	−1	3	−15
(4) $(2)^3$ $(1)^2$	−6	45	−420	.	.	.	.	.	.	.	.	−3	30
(4) $(2)^2$ $(1)^4$	1	−15	210	.	.	.	.	.	.	.	.	.	−3
(4) (2) $(1)^6$	15	1	−28	.	.	.	.	.	.	.	.	.	.
(4) $(1)^8$	210	28	1	.	.	.	.	.	.	.	.	.	.
$(3)^4$	.	.	.	1	−1	2	.	2	−8	40	.	−6	40
$(3)^3$ (2) (1)	.	.	.	1	1	−3	.	−4	20	−120	.	18	−140
$(3)^3$ $(1)^3$	.	.	.	1	3	1	.	.	−4	40	.	.	20
$(3)^2$ $(2)^2$	.	.	.	.	.	.	1	−1	3	−15	−4	11	−80
$(3)^2$ (2) $(1)^3$	.	.	.	2	4	.	1	1	−6	45	.	−9	80
$(3)^2$ $(1)^6$	.	.	.	4	16	4	3	6	1	−15	.	.	−20
(3) $(2)^3$ (1)	.	.	.	10	60	20	15	45	15	1	.	.	.
(3) $(2)^2$ $(1)^3$	.	.	.	.	.	.	4	.	.	.	1	−3	15
(3) $(2)^2$ $(1)^3$	.	.	.	6	18	.	10	9	.	.	3	1	−10
(3) (2) $(1)^5$	5	.	.	20	100	20	40	70	10	.	15	10	1
(3) $(1)^9$	105	7	.	70	490	140	210	525	140	7	105	105	21
	1890	252	9	280	2520	840	1260	3780	1260	84	945	1260	378
$(2)^6$	.	.	.	.	.	.	.	.	.	.	.	.	.
$(2)^4$ $(1)^4$	.	.	.	.	.	.	20	.	.	.	10	.	.
$(2)^4$ $(1)^4$	6	.	.	24	96	.	88	72	.	.	52	16	.
$(2)^3$ $(1)^6$	90	3	.	120	720	120	460	720	90	.	330	200	18
$(2)^2$ $(1)^8$	1120	84	1	560	4480	1120	2800	6160	1400	56	2520	2240	392
(2) $(1)^{10}$	12600	1470	45	2800	28000	8400	18900	50400	14700	840	22050	25200	6300
$(1)^{12}$	103950	13860	495	15400	184800	61600	138600	415800	138600	9240	207900	277200	83160

Table 1.12 (*cont.*)

w = 12 (vi)	$[321^3]$	$[31^5]$	$[2^6]$	$[2^421^3]$	$[2^41^6]$	$[2^31^9]$	$[2^21^{12}]$	$[21^{15}]$	$[1^{18}]$
(12)	40320	−362880	−120	720	−5040	40320	−362880	3628800	−39916800
(11) (1)	−35280	362880	.	−240	2880	−30240	322560	−3628800	43545600
(10) (2)	−20160	181440	144	−624	3600	−25920	221760	−2177280	23950080
(10) $(1)^2$	15120	−181440	.	24	−720	10800	−141120	1814400	−23950080
(9) (3)	−18480	161280	.	−240	2112	−17760	161280	−1612800	17740800
(9) (2) (1)	17640	−181440	.	240	−2208	20160	−201600	2217600	−26611200
(9) $(1)^3$	−4200	60480	.	.	96	−2400	40320	−604800	8870400
(8) (4)	−15120	136080	90	−300	1908	−15120	136080	−1360800	14968800
(8) (3) (1)	16800	−166320	.	60	−1200	13680	−147840	1663200	−19958400
(8) $(2)^2$	5040	−45360	−90	270	−1314	8640	−70560	680400	−7484400
(8) (2) $(1)^2$	−7560	90720	.	−30	612	−7560	90720	−1134000	14968800
(8) $(1)^4$	840	−15120	.	.	−6	360	−8400	151200	−2494800
(7) (5)	−13824	124416	.	−240	1728	−13824	124416	−1244160	13685760
(7) (4) (1)	13860	−142560	.	120	−1152	11880	−126720	1425600	−17107200
(7) (3) (2)	9960	−86400	.	240	−1632	12240	−105600	1036800	−11404800
(7) (3) $(1)^2$	−7560	86400	.	.	288	−5040	67200	−864000	11404800
(7) $(2)^2$ (1)	−4410	45360	.	−120	864	−7020	66240	−712800	8553600
(7) (2) $(1)^3$	2100	−30240	.	.	−96	1800	−26880	388800	−5702400
(7) $(1)^5$	−126	3024	.	.	.	−36	1344	−30240	570240
(6)^3	−6720	60480	40	−120	840	−6720	60480	−604800	6652800
(6) (5) (1)	12936	−133056	.	80	−1056	11088	−118272	1330560	−15966720
(6) (4)(2)	8400	−75600	−120	280	−1512	10800	−92400	907200	−9979200
(6) (4) $(1)^2$	−6300	75600	.	−20	312	−4500	58800	−756000	9979200
(6) $(3)^2$	4760	−40320	.	40	−496	4400	−40320	403200	−4435200
(6) (3) (2) (1)	−9240	90720	.	−80	1024	−9960	100800	−1108800	13305600
(6) (3) $(1)^3$	2240	−30240	.	.	−32	1160	−20160	302400	−4435200
(6) $(2)^3$	−840	7560	40	−80	336	−2040	15960	−151200	1663200
(6) $(2)^2$ $(1)^2$	1890	−22680	.	20	−264	2790	−31080	378000	−4989600
(6) (2) $(1)^4$	−420	5560	.	.	8	−300	5880	−100800	1663200
(6) $(1)^6$	14	−504	.	.	.	2	−168	5040	−110880
(5)^3 (2)	4032	−36288	.	120	−720	5184	−44352	435456	−4790016
(5)^3 $(1)^3$	−3024	36288	.	.	144	−2160	28224	−362880	4790016
(5) (4) (3)	8316	−72576	.	120	−960	7992	−72576	725760	−7983360
(5) (4) (2) (1)	−7938	81648	.	−120	1008	−9072	90720	−907920	11975040
(5) (4) $(1)^3$	1890	−27216	.	.	−48	1080	−18144	272160	−3991680
(5) $(3)^2$ (1)	−4648	44352	.	.	288	−3600	39424	−443520	5322240
(5) (3) $(2)^2$	2814	−24192	.	−120	672	−4572	37632	−362880	3991680
(5) (3) (2) $(1)^3$	−4284	48384	.	.	−288	3960	−48384	604800	−7983360
(5) (3) $(1)^5$	490	−8064	.	.	.	−180	4480	−80640	1330560
(5) $(2)^3$ (1)	−735	7560	.	40	−240	1764	−15792	166320	−1995840
(5) $(2)^3$ $(1)^3$	525	−7560	.	.	48	−720	9744	−136080	1995840
(5) (2) $(1)^5$	−63	1512	.	.	.	36	−1008	21168	−399168
(5) $(1)^7$	1	−72	.	.	.	.	16	−720	19008
(4)^3	−1260	11340	−15	30	−162	1260	−11340	113400	−1247400
(4)^3 (3) (1)	4200	−41580	45	−30	312	−3420	36960	−415800	4989600
(4)^3 $(2)^2$	−1260	11340	.	−75	333	−2160	17640	−170100	1871100
(4)^3 (2) $(1)^2$	1890	−22680	.	15	−162	1890	−22680	283500	−3742200
(4)^3 $(1)^6$	−210	3780	.	.	3	−90	2100	−37800	623700
(4) (3)^2 (2)	−3010	25200	.	−60	456	−3540	30800	−302400	3326400
(4) (3)^2 $(1)^3$	2310	−25200	.	.	−72	1440	−19600	252000	−3326400
(4) (3) $(2)^3$ (1)	2730	−26460	.	60	−480	4050	−38640	415800	−4989600
(4) (3) (2) $(1)^3$	−1330	17640	.	.	48	−1020	15680	−226800	3326400
(4) (3) $(1)^5$	84	−1764	.	.	.	18	−784	17640	−332640
(4) $(2)^4$	105	−945	−15	20	−72	405	−3045	28350	−311850
(4) $(2)^3$ $(1)^2$	−315	3780	.	−10	84	−765	7980	−94500	1247400
(4) $(2)^2$ $(1)^4$	105	−1890	.	.	−6	135	−2310	37800	−623700
(4) (2) $(1)^6$	−7	252	.	.	.	−3	140	−3780	83160
(4) $(1)^8$	.	−9	.	.	.	.	−1	90	−2970
(3)^4	−280	2240	.	.	24	−240	2240	−22400	246400
(3)^3 (2) (1)	1120	−10080	.	.	−96	1080	−11200	123200	−1478400
(3)^3 $(1)^3$	−280	3360	.	.	.	−120	2240	−33600	492800
(3)^3 $(2)^2$	315	−2520	.	20	−104	670	−5320	50400	−554400
(3)^3 (2) $(1)^3$	−735	7560	.	.	72	−900	10360	−126000	1663200
(3)^3 $(1)^6$	175	−2520	.	.	.	90	−1960	33600	−554400
(3) $(2)^4$ (1)	−7	168	.	.	.	.	56	−1680	36960
(3) $(2)^3$ $(1)^3$	−105	945	.	−10	56	−390	3360	−34650	415800
(3) $(2)^2$ $(1)^5$	105	−1260	.	.	−16	220	−2800	37800	−554400
(3) (2) $(1)^7$	−21	378	.	.	.	−18	448	−8820	166320
(3) $(1)^9$	1	−36	.	.	.	.	−16	600	−15840
36		1	.	.	.	.	.	−10	440
$(2)^6$	.	.	1	−1	3	−15	105	−945	10395
$(2)^5$ $(1)^2$	.	.	1	1	−6	45	−420	4725	−62370
$(2)^4$ $(1)^4$	.	.	3	6	1	−15	210	−3150	51975
$(2)^3$ $(1)^6$	.	.	15	45	15	1	−28	630	−13860
(2) $(1)^{10}$	16	.	105	420	210	28	1	−45	1485
$(1)^{12}$	480	10	945	4725	3150	630	45	1	−66
	7920	220	10395	62370	51975	13860	1485	66	1

MISCELLANEA

On the efficiency of the method of moments and Neyman's type A distribution

By L. R. SHENTON

In a paper by Neyman (1939) several types of compound Poisson distributions are derived and their application to empirical data mentioned. Neyman remarks that the method of fitting requires investigation. It is the object of this note to consider the efficiency of the method of moments and Neyman's type A distribution of two parameters.

The efficiency of the method of moments, with particular reference to parameters of scale and location, has been discussed by R. A. Fisher (1921), in connexion with the Pearson system of frequency curves. More recently Fisher (1941) has used the covariance and information matrices to find an expression for the efficiency of the method of moments applied to the negative binomial distribution. The chief difficulty appears to be the evaluation of the information determinant, and the process given here may have applications in other cases.

1. Neyman's type A distribution

The probability function is given by

$$\exp\{-m_1[1-e^{m_2(t-1)}]\} = \sum_{x=0}^{\infty} P_x t^x, \tag{1}$$

with

$$P_x = \frac{e^{-m_1} m_2{}^x}{x!}\left\{0^x + \frac{\lambda' 1^x}{1!} + \frac{\lambda' 2 2^x}{2!} + \frac{\lambda' 3 3^x}{3!} + \ldots\right\}, \tag{2}$$

where

$$\lambda' = m_1 e^{-m_2}, \quad 0^x = 1, \quad x = 0,$$
$$= 0, \quad x \neq 0.$$

The cumulants K_r are given by

$$\sum_{r=1}^{\infty} K_r t^r/r! = \sum_{s=1}^{\infty} m_1 m_2^s (e^t-1)^s/s!. \tag{3}$$

The first few are recorded for later use

$$\left.\begin{aligned} K_1 &= m_1 m_2, \qquad K_2 = m_1 m_2 (1+m_2), \\ K_3 &= m_1 m_2 (1+3m_2+m_2^2), \\ K_4 &= m_1 m_2 (1+7m_2+6m_2^2+m_2^3), \\ K_5 &= m_1 m_2 (1+15m_2+25m_2^2+10m_2^3+m_2^4), \\ K_6 &= m_1 m_2 (1+31m_2+90m_2^2+65m_2^3+15m_2^4+m_2^5), \end{aligned}\right\} \tag{4}$$

The relations

$$K_{r+1} = m_2\left\{K_r + rK_{r-1} + \frac{r(r-1)}{2} K_{r-2} + \ldots + rK_1 + m_1\right\} \tag{5}$$

and

$$K_{r+1} = m_2\left\{K_r + \frac{\partial K_r}{\partial m_2}\right\},$$

are easily proved from (3), the second expression being useful if high-order cumulants are required.

By the method of moments m_1 and m_2 are estimated from

$$\bar{\lambda}_1 = m_1 m_2, \quad \bar{\lambda}_2 = m_1 m_2 (1+m_2),$$

where $\bar{\lambda}_1$ and $\bar{\lambda}_2$ are the first two sample moments about the mean. For large samples of n, we find

$$n \operatorname{var} m_1 = m_1\{2 + m_2^2 + 2m_1(1+m_2)^2\}/m_2^2,$$
$$n \operatorname{var} m_2 = \{2 + m_2 + 2m_1(1+m_2)^2\}/m_1,$$
$$n \operatorname{cov}(m_1, m_2) = -2\{1 + m_1(1+m_2)^2\}/m_2,$$

so that the covariance matrix

$$\begin{bmatrix} \operatorname{var} m_1 & \operatorname{cov}(m_1 m_2) \\ \operatorname{cov}(m_1, m_2) & \operatorname{var} m_2 \end{bmatrix}$$

has a determinant of value

$$\{1 + (1+m_2)^2 + 2m_1(1+m_2)^2\}/n^2 m_2. \tag{6}$$

To find the efficiency of fitting the first two moments, we require the determinant of the information matrix.

2. Likelihood equations and information matrix

By differentiating the G.F. (1) we find

$$\frac{\partial P_x}{\partial m_2} = \{xP_x - (x+1)P_{x+1}\}/m_2 \quad \text{and} \quad \frac{\partial P_x}{\partial m_1} = -P_x + \frac{(x+1)}{m_1 m_2} P_{x+1}. \tag{7}$$

The likelihood of the sample $(n_0, n_1, n_2, \ldots, n_n)$ is

$$L = \prod_{x=0}^{n} (P_x)^{n_x},$$

and so for optimum statistics $\hat{m}_1$, $\hat{m}_2$

$$\Sigma' n_x (x+1) \frac{P_{x+1}}{P_x} = n\hat{m}_1\hat{m}_2 = n\bar{\lambda}_1,$$

with Σ' indicating summation over the sample. $m_1 m_2$ is therefore efficiently estimated by the mean, and to complete the solution we must approximate to

$$\Sigma' n_x (x+1) \frac{P_{x+1}}{P_x} = n\bar{\lambda}_1. \tag{8}$$

If m_1 and m_2 are estimates by moments, they can be improved by (8). For $m_1 m_2 = \hat{m}_1\hat{m}_2 = \bar{\lambda}_1$, and writing $P_x(m_2)$ to indicate P_x with $m_1 = \bar{\lambda}_1/m_2$, we have

$$\Sigma n_x (x+1) \frac{\partial}{\partial m_2} \frac{P_{x+1}(m_2)}{P_x(m_2)} = \frac{1}{m_2} \Sigma' n_x (x+1) \frac{P_{x+1}}{P_x} - \frac{(m_2+1)}{m_2^2} \Sigma' \left\{ \frac{(x+1)(x+2)P_{x+2}}{P_x} - \frac{(x+1)^2 P_{x+1}}{P_x^2} \right\}. \tag{9}$$

Thus if $$F(\hat{m}_2) = \Sigma' n_x (x+1) \frac{P_{x+1}}{P_x} - n\bar{\lambda}_1 \quad \text{and} \quad \hat{m}_2 = m_2 + \delta m_2$$

then $$\delta m_2 = -F(m_2)/F'(m_2),$$

where (9) gives $F'(m_2)$. The improved solution is therefore obtained by the use of frequencies calculated in the moment solution.

We now proceed to consider the efficiency. Using equations (7) and (4) we have

$$E\left(-\frac{\partial^2}{\partial m_1^2} \log P_x\right) = E\left(\frac{1}{P_x}\frac{\partial P_x}{\partial m_1}\right)^2 = -1 + \phi/m_1^2 m_2^2 = i_{m_1 m_1},$$

$$E\left(-\frac{\partial^2}{\partial m_1 \partial m_2} \log P_x\right) = E\left(\frac{1}{P_x^2}\frac{\partial P_x}{\partial m_1}\frac{\partial P_x}{\partial m_2}\right) = 1 + m_1 - \phi/m_1 m_2^2 = i_{m_1 m_2},$$

$$E\left(-\frac{\partial^2}{\partial m_2^2} \log P_x\right) = E\left(\frac{1}{P_x}\frac{\partial P_x}{\partial m_2}\right)^2 = \frac{m_1}{m_2} - m_1(1+m_1) + \phi/m_2^2 = i_{m_2 m_2},$$

where $$\phi = E\{(x+1)^2 P_{x+1}^2/P_x^2\},$$

and for the information determinant

$$\begin{vmatrix} ni_{m_1 m_1} & ni_{m_1 m_2} \\ ni_{m_1 m_2} & ni_{m_2 m_2} \end{vmatrix} = n^2\{(1+m_2)\phi - m_1 m_2^2(m_1 + m_1 m_2 + m_2)\}/m_1 m_2^2. \tag{10}$$

3. The evaluation of $\phi = E\{(x+1)^2 P_{x+1}^2/P_x^2\}$

Since

$$P_x = e^{-m_1}\frac{m_2^x}{x!}\left\{0^x + \frac{\lambda' 1^x}{1!} + \frac{\lambda'^2 2^x}{2!} + \frac{\lambda'^3 3^x}{3!} + \ldots\right\}$$

$$= e^{-m_1}\frac{m_2^x}{x!}\{A_0 + A_1 x + A_2 x(x-1) + A_3 x(x-1)(x-2) + \ldots\},$$

where the A's may be determined by the Gregory-Newton formula of interpolation, it is evident that we may set up orthogonal polynomials with respect to P_x, defined by

$$\sum_0^\infty \theta_r(x)\theta_s(x) P_x = 0, \quad r \neq s,$$

$$\neq 0, \quad r = s,$$

with $\theta_0 = 1$, say. With these polynomials we may then find an expression for $(x+1)\,P_{x+1}$ in the form

$$\{B_0 + B_1\theta_1(x) + B_2\theta_2(x) + \ldots\}\,P_x, \tag{11}$$

where the B's are functions of m_1 and m_2. In fact we have

$$\Sigma(x+1)\,P_{x+1}\theta_r(x) = B_r\Sigma\theta_r^2(x)\,P_x.$$

Squaring (11) and summing, we have

$$\phi = \sum_0^\infty (x+1)^2\frac{P_{x+1}^2}{P_x} = B_0^2\Sigma P_x + B_1^2\Sigma\theta_1^2(x)\,P_x + B_2^2\Sigma\theta_2^2(x)\,P_x + B_3^2\Sigma\theta_3^2(x)\,P_x + \ldots. \tag{12}$$

This is a series of positive terms, the first two of which amount to $m_1m_2^2[m_1(1+m_2)+m_2]/(1+m_2)$, and thus our expression (10) for the determinant of the information matrix is a series of positive terms. The first term of significance turns out to be the fourth in (12). The value of (12) can therefore be found in two stages; (*a*) the determination of the orthogonal polynomials, (*b*) the evaluation of the B's.

(*a*) The evaluation of the θ's will be illustrated by finding $\theta_3(x)$. It appears most convenient to assume

$$\theta_3(x) = (x-\lambda_1)^3 - \mu_2 + A[(x-\lambda_1)^2 - \mu_2] + B[x-\lambda_1],$$

where λ_1 is m_1m_2 and the μ's refer to the moments of P_x about the mean. The orthogonality conditions lead to

$$\left.\begin{aligned} -\theta_3(x) + (x-\lambda_1)^3 - \mu_2 + A[(x-\lambda_1)^2 - \mu_2] + B(x-\lambda_1) &= 0,\\ \mu_4 \qquad\quad + A\mu_2 \qquad\quad + B\mu_2 &= 0,\\ \mu_5 - \mu_2\mu_2 + A[\mu_4 - \mu_2^2] \quad + B\mu_3 &= 0,\\ -\Sigma\theta_3^2(x)\,P_x + \mu_6 - \mu_3^2 \quad + A[\mu_5 - \mu_2\mu_2] \quad + B\mu_4 &= 0. \end{aligned}\right\} \tag{13}$$

The moments in (13) may be replaced by cumulants, using well-known formulae

$$K_2 = \mu_2, \quad K_3 = \mu_3, \quad K_4 + 3K_2^2 = \mu_4,$$
$$K_5 + 10K_3K_2 = \mu_5, \quad K_6 + 15K_4K_2 + 10K_3^2 + 15K_2^3 = \mu_6,$$

etc., and these in turn expressed in terms of m_1 and m_2 by means of (4).

(*b*) For B_3 we have $\quad B_3\sum_0^\infty \theta_3^2(x)\,P_x = \sum_0^\infty (x+1)\,P_{x+1}\theta_3(x),$

so that we require the value of $\quad \sum_0^\infty (x-\lambda_1+\lambda_1)\,\theta_3(x-1)\,P_x.$

Expanding and simplifying, this becomes

$$\mu_4 - 3\mu_3 + 3(1-\lambda_1)\mu_2 - \lambda_1 + A[\mu_2 - 2\mu_2 + \lambda_1] + B[\mu_2 - \lambda_1].$$

Hence from (13) we have

$$B_3^2\sum_0^\infty \theta_3^2(x)\,P_x = \frac{\begin{vmatrix} -3\mu_3 + 3(1-\lambda_1)\mu_2 - \lambda_1 & -2\mu_2 + \lambda_1 & -\lambda_1 \\ \mu_4 & \mu_2 & \mu_2 \\ \mu_5 - \mu_2\mu_2 & \mu_4 - \mu_2^2 & \mu_3 \end{vmatrix}^2}{\begin{vmatrix} \mu_2 & \mu_2 \\ \mu_4 - \mu_2^2 & \mu_3 \end{vmatrix}\begin{vmatrix} \mu_4 & \mu_2 & \mu_2 \\ \mu_5 - \mu_2\mu_2 & \mu_4 - \mu_2^2 & \mu_3 \\ \mu_6 - \mu_3^2 & \mu_5 - \mu_2\mu_2 & \mu_4 \end{vmatrix}}$$

$$= \frac{4m_1m_2^2[1+m_1(1+m_2)(4+m_2)]^2}{[2+2m_2+m_2^2+2m_1(1+m_2)^3]\alpha},$$

where

$$\alpha = 12 + 12m_2 + 6m_2^2 + 2m_2^3 + m_1(48 + 144m_2 + 144m_2^2 + 80m_2^3 + 25m_2^4 + 2m_2^5) + 24m_1^2(1+m_2)^3(2+2m_2+m_2^2) + 12m_1^3(1+m_2)^6. \tag{14}$$

For the earlier polynomials it may be verified that

$$\theta_0(x) = 1, \quad \Sigma\theta_0^2(x)\,P_x = 1, \quad B_0^2 = m_1^2m_2^2,$$
$$\theta_1(x) = x - m_1m_2, \quad B_1^2\Sigma\theta_1^2(x)\,P_x = m_1m_2^3(1+m_2),$$
$$\theta_2(x) = (x-m_1m_2)^2 - \frac{(1+3m_2+m_2^2)}{1+m_2}(x-m_1m_2) - m_1m_2(1+m_2),$$
$$B_2^2\Sigma\theta_2^2(x)\,P_x = \frac{m_1m_2^4}{(1+m_2)[1+(1+m_2)^2+2m_1(1+m_2)^3]}.$$

Inserting these values in (10) and multiplying by the covariance determinant (6), we have for the efficiency

$$\frac{1}{E} > 1 + \frac{4m_2(1+m_2)[1+m_1(1+m_2)(4+m_2)]^2}{\alpha},$$

where α is given by (14).

Table 1 gives an upper bound to the value of percentage efficiency, E, for various values of m_1 and m_2. It is clear from Table 1 that if m_2 is small, say less than 0·2, then E is generally in the region of 90 % or above. For $0{\cdot}2 < m_2 < 1{\cdot}0$, E may be as low as 70 %, but otherwise it lies between 75 and 90 %, and for these values it is not easy to decide whether the 'improved' fit is worth the additional computation involved. If $m_2 \geqslant 1{\cdot}0$ and $m_1 < 3{\cdot}0$, E may be less than 70 % and in this case the maximum likelihood approximation (9) can be applied.

Table 1

m_2 \ m_1	0·1	0·5	1·0	2·0	3·0	4·0	6·0	10·0
0·1	96	93	93	93	94	95	96	97
0·2	92	88	87	88	89	91	93	95
0·5	82	76	77	80	83	85	89	92
1·0	73	67	69	76	80	83	87	91
2·0	65	61	67	75	81	84	88	92
3·0	62	60	67	77	82	85	89	93
5·0	59	59	68	79	84	87	91	94

4. A numerical illustration

That Neyman's type A distribution of two parameters and the negative binomial are related, in that they may arise as the result of heterogeneity in the population, has been pointed out by W. Feller (1943). The similarity of the two may be appreciated by comparing the first three moments:

	Neyman's type A parameters m_1 and m_2	Negative binomial parameters m_1 and m_2
μ_1	$m_1 m_2$	$m_1 m_2$
μ_2	$m_1 m_2(1+m_2)$	$m_1 m_2(1+m_2)$
μ_3	$m_1 m_2(1+3m_2+m_2^2)$	$m_1 m_2(1+3m_2+2m_2^2)$

It is thus reasonable to suppose that data satisfactorily described by the negative binomial may be suitable for our present purpose provided the parameters satisfy the conditions $m_2 \geqslant 1{\cdot}0$, $m_1 < 3{\cdot}0$. Such a distribution is given by Ove Lundberg (1940) concerning insurance claims and incapacity caused by sickness or accident.

The moments of the distribution turn out to be

$$\bar{\lambda}_1 = 2{\cdot}805{,}871, \quad \bar{\lambda}_2 = 6{\cdot}404{,}549,$$

from which, using the method of moments, we obtain the estimates

$$m_1 = 2{\cdot}187{,}723, \quad m_2 = 1{\cdot}282{,}553.$$

Frequencies for these values are obtained by using the expression due to Beall (Neyman, 1939), namely,

$$(x+1)\,P_{x+1} = m_1 m_2 e^{-m_2}\left\{P_x + \frac{m_2}{1!}P_{x-1} + \frac{m_2^2}{2!}P_{x-2} + \ldots\right\},$$

with $m_1 m_2 e^{-m_2} = 0{\cdot}778{,}1466$, $P_0 = e^{-a}$, where $a = m_1(1-e^{-m_2})$,

so that $P_0 = 0{\cdot}205{,}7682$.

Table 2

1	2	3	4	5	6
x	Observed frequency	P_x by moments	P_x by likelihood	χ^2 (moments)	χ^2 (likelihood)
0	187	217·29	197·15	4·22	0·52
1	185	169·08	178·00	1·50	0·28
2	200	174·22	181·27	3·81	1·94
3	164	147·79	153·44	1·78	0·73
4	107	114·14	117·67	0·45	0·97
5	68	82·63	83·97	2·59	3·04
6	49	56·70	56·49	1·05	0·99
7	39	37·17	36·18	0·09	0·22
8	21	23·44	22·21	0·25	0·07
9	12	14·29	13·15	0·37	0·10
10	11	8·45	7·54	0·77	1·59
11	2	4·87	4·20	0·44	1·83
12	5	2·74	2·28	—	—
13	2 (13)	1·51 (10·81)	1·21 (8·95)	—	—
14	3	0·81	0·63	—	—
15	1	0·43	0·32	—	—
⩾16	—	0·45	0·31	—	—
	1056	—	—	17·32	12·28
	$\log_{10} L$	−985·948,937	−984·902,555	—	—

The fit by moments is given in column 3 of Table 2. For an improved fit we set up a table of values consisting of (*a*) $(x+1)\,P_{x+1}/P_x$, (*b*) $(x+1)^2\,P^2_{x+1}/P^2_x$, which is easily found from (*a*), (*c*) $(x+1)(x+2)\,P_{x+2}/P_x$, in each of which x takes the values 0, 1, 2, ..., to 15. In the present case we have

$$\Sigma n_x(x+1)\,P_{x+1}/P_x = 2983{\cdot}569{,}210,$$
$$\Sigma n_x(x+1)^2\,P^2_{x+1}/P^2_x = 10550{\cdot}201{,}104,$$
$$\Sigma n_x(x+1)(x+2)\,P_{x+2}/P_x = 12126{\cdot}916{,}881.$$

Using these in (9) we find $\delta m_2 = -0{\cdot}14863$, and the improved estimates $\hat{m}_1 = 2{\cdot}474{,}481$, $\hat{m}_2 = 1{\cdot}133{,}923$. Since with these values $\Sigma n_x(x+1)\,P_{x+1}/P_x = 2965{\cdot}031{,}424$, it is clear we are much nearer a maximum likelihood solution for which

$$\Sigma n_x(x+1)\,P_{x+1}/P_x = n\bar{\lambda}_1 = 2963.$$

The fit with the improved estimates is shown in column 4 of the table. The improvement is probably exaggerated by the value of χ^2, where we have grouped the last five frequencies. A more reliable idea is given by $\log_{10} L$, shown at the bottom of the table, and it is evident that an improvement has been effected.

REFERENCES

Feller, W. (1943). On a general class of 'contagious' distributions. *Ann. Math. Statist.* **14**, 389.

Fisher, R. A. (1921). On the mathematical foundations of theoretical statistics. *Philos. Trans.* A, **222**, 309.

Fisher, R. A. (1941). The negative binomial distribution. *Ann. Eugen., Lond.*, **11**, 182.

Lundberg, O. (1940). *On Random Processes and their Application to Sickness and Accident Statistics.* Uppsala: Almquist and Wiksells.

Neyman, J. (1939). On a new class of 'contagious' distributions applicable in entomology and bacteriology. *Ann. Math. Statist.* **10**, 35.

Large-sample theory of sequential estimation

By F. J. ANSCOMBE

Haldane (1945) and Finney (1949) have considered a sequential method of sampling a binomial population, called inverse sampling, in which sampling terminates when a specified number k of individuals have been found possessing the attribute. If the proportion θ of individuals with the attribute in the population is small, θ is estimated with coefficient of variation depending only on the value of k. Thus the coefficient of variation of the estimate can be specified in advance. Tweedie (1945) has considered inverse sampling more generally.

Several recent papers in the *Annals of Mathematical Statistics* have dealt with the estimation of parameters in sequential sampling; in particular, Blackwell & Girshick (1947) have given a lower bound to the variance of an unbiased estimate based on a sufficient statistic. A general problem of sequential estimation suggests itself, namely, to formulate a rule of sampling such that an unknown population parameter can be estimated with specified accuracy and with minimum expected sample size, whatever the true value of the parameter. The accuracy of estimation might be specified by width of confidence interval, or by variance or coefficient of variation of the estimate, or in any other such way. It is the object of this note to point out that, while an exact small-sample solution of the problem is in general very difficult, the large-sample theory (valid for large expected sample size and high accuracy of estimation) is relatively easy.

Let us consider the estimation of the mean θ of a population of which the variance is a known finite function $v(\theta)$ of θ. The observations will be denoted by $z_1, z_2, \ldots$, and the cumulative sum of the first m observations by Z_m. Sampling continues until a certain inequality is satisfied, of the form $Z_m \geqslant k(m)$ or $Z_m \leqslant k(m)$, where $k(m)$ is a function of m determined in advance. The value of m for which the inequality is first satisfied will be denoted by n. If we represent the sampling on a diagram in which m is abscissa and the ordinate is denoted generally by y, n is abscissa of the point where the sample path $y = Z_m$ first meets or crosses the boundary $y = k(m)$. We shall estimate θ by

$$\hat{\theta} = \frac{k(n)}{n}. \tag{1}$$

It will be shown that, under certain conditions, $\hat{\theta}$ is asymptotically normally distributed with mean θ and variance $v(\theta)/n_0$, where n_0 satisfies

$$k(n_0) = \theta n_0, \tag{2}$$

i.e. n_0 is abscissa of the point where the mean path $y = \theta m$ intersects the boundary. θ will therefore be estimated with specified variance a^2 if the equation of the boundary is

$$\frac{1}{m} v\left(\frac{y}{m}\right) = a^2, \tag{3}$$

and with specified coefficient of variation b if the boundary is

$$\frac{m}{y^2} v\left(\frac{y}{m}\right) = b^2. \tag{4}$$

To establish this result (a sketch proof only will be given), we consider a sequence of sequential samplings performed on the same population (so that θ is contant) and defined by a sequence of boundaries $y = k(m)$ such that $n_0 \to \infty$. By (2) we have $|k(n_0)| \to \infty$ also, if $\theta \neq 0$. Let us suppose that the boundary system satisfies the following conditions:

(i) The boundary is approximately linear in the neighbourhood of n_0. More exactly, if m is confined to an interval $(n_0 - c\sqrt{n_0}, n_0 + c\sqrt{n_0})$, where c is constant, then as $n_0 \to \infty$

$$k(m) = k(n_0) + (m - n_0) k'(n_0) + O(1). \tag{5}$$

(ii) We can ignore the possibility that the boundary will be crossed elsewhere than in the neighbourhood of n_0, i.e. by choosing c sufficiently large we can make the chance arbitrarily small that $|n - n_0| > c\sqrt{n_0}$, if n_0 is large.

(iii) The boundary crosses the mean path $y = \theta m$ at a non-zero angle. More exactly, $k'(n_0)$ tends to a limit not equal to θ as $n_0 \to \infty$.

It has been assumed here that the limit of $k'(n_0)$ is finite. If it is infinite, condition (i) must be reformulated: if $k(m)$ is confined to an interval $(k(n_0) - c\sqrt{k(n_0)},\ k(n_0) + c\sqrt{k(n_0)})$, where c is constant, then as $n_0 \to \infty$

$$m = n_0 + O(1). \tag{6}$$

Conditions (i) and (iii) are in fact satisfied by the boundaries defined by equations (3) and (4). For the first of these,

$$n_0 = \frac{v(\theta)}{a^2}, \tag{7}$$

$$k'(n_0) = \theta + \frac{v(\theta)}{v'(\theta)}; \tag{8}$$

and if we consider a sequence of boundaries for which $a \to 0$, we have $n_0 \to \infty$, $k'(n_0)$ is a constant (finite or infinite) not equal to θ if $v(\theta) > 0$, and (provided $k'(n_0)$ is finite) $k''(n_0) = O(n_0^{-1})$. Similarly for the second, when $\theta \neq 0$,

$$n_0 = \frac{v(\theta)}{b^2\theta^2}, \tag{9}$$

$$k'(n_0) = \frac{\theta v'(\theta) - v(\theta)}{\theta v'(\theta) - 2v(\theta)}\theta, \tag{10}$$

and the same conclusions follow. Condition (ii), however, is not always satisfied by these boundaries, and they may need modification, as described below.

Assuming that all three conditions are satisfied, we consider first the case $\lim k'(n_0)$ finite, and show that if m has any value in the range $|m - n_0| < c\sqrt{n_0}$,

$$\text{prob}\,(n \leqslant m) = \text{prob}\,(Z_m \geqslant k(m)) + o(1) \tag{11}$$

as $n_0 \to \infty$ if $k'(n_0) < \theta$, while the inequality on the right-hand side is reversed if $k'(n_0) > \theta$. Here Z_m is the cumulative sum of m observations, taken without regard to whether the boundary has been crossed or not. Let us suppose that $k'(n_0) < \theta$. Then the left-hand side of (11) exceeds the first member on the right-hand side by the chance that a path reaches the boundary for some $n < m$ and then recrosses it to give $Z_m < k(m)$. This requires that $\sum_{i=n+1}^{m} z_i$ shall be less than $(m-n)\,k'(n_0)$, while its mean is $(m-n)\,\theta$ and its variance is $(m-n)\,v(\theta)$. Now as $n_0 \to \infty$ the variance of n is of order n_0, i.e. tends to a finite non-zero multiple of n_0. Hence, applying the central limit theorem, and integrating over values of $n < m$, we see that the chance in question is of order $\Phi(-m^{\frac{1}{2}})$, where Φ denotes the normal integral, and so is $o(1)$. Similarly if $k'(n_0) > \theta$. It may be noted in passing that if $z_i \geqslant 0$ always and if $k'(m) \leqslant 0$ for all m, the boundary once reached cannot be recrossed, and equation (11), without the $o(1)$ on the right-hand side, gives the exact distribution of n in terms of the distribution of Z_m.

Now the distribution of Z_m is asymptotically normal with mean θm and variance $v(\theta)\,m$. Let us set $m = n_0 + \mu$, where $|\mu| < c\sqrt{n_0}$. Then asymptotically

$$\text{prob}\,(Z_m \geqslant k(m)) = \text{prob}\,(Z_m \geqslant k(n_0) + \mu k'(n_0)) = \Phi\left(\frac{(\theta - k'(n_0))\,\mu}{[(n_0+\mu)\,v(\theta)]^{\frac{1}{2}}}\right). \tag{12}$$

This takes values other than 0 and 1 only when $\mu = O(\sqrt{n_0})$, and so we may neglect the μ in the denominator. Thus n is asymptotically normally distributed with mean n_0 and variance $n_0 v(\theta)\,[\theta - k'(n_0)]^{-2}$. But from (1) we have asymptotically

$$\hat{\theta} = \frac{k(n_0) + \mu k'(n_0)}{n_0 + \mu} = \theta - \frac{\mu}{n_0}(\theta - k'(n_0)). \tag{13}$$

Hence we have the desired result that $\hat{\theta}$ is asymptotically normally distributed with mean θ and variance $v(\theta)/n_0$.

If the boundary is wholly vertical, the result is still true, since the sample size n is now fixed. It can also be established if $\lim k'(n_0) = \infty$ for some value of θ, using equation (6).

We can replace n_0 by $\mathrm{E}(n)$ in the variance. The asymptotic variance is then equal to Blackwell & Girshick's lower bound to the variance of an unbiased estimate of θ. We may also replace $k(n)$ by Z_n in the numerator of (1).

Let us now consider some examples of boundaries. First suppose $v(\theta)$ is a known constant. Then equation (3) represents a vertical boundary, i.e. fixed sample size, with

$$m = v/a^2. \tag{14}$$

Stein & Wald (1947) have demonstrated the optimum character of this boundary. Equation (4) gives a parabolic boundary,

$$y^2 = (v/b^2)\, m. \tag{15}$$

This boundary, as it stands, passes through the origin. The above theory will apply if we can ignore the possibility that the boundary will be reached near the origin instead of near $m = n_0$. The boundary should therefore be modified for m small so that it does not approach very close to the origin. If $|\theta|$ is not too large θ will be estimated with the specified coefficient of variation b, but it will be estimated with greater accuracy than required if $|\theta|$ is so large that the sample path is likely to meet the modified part of the boundary.

Next, let us consider the Poisson distribution, for which $v(\theta) = \theta$. Equation (3) gives a parabolic boundary,

$$y = a^2 m^2, \tag{16}$$

which again needs modification for m small (so that if θ is near 0 it will be estimated with lower variance than a^2). Equation (4) gives a horizontal boundary,

$$y = 1/b^2. \tag{17}$$

For the type III distribution of a variance estimate of ν degrees of freedom from a normal population with variance θ, we have $v(\theta) = (2/\nu)\,\theta^2$. Equation (3) gives the curve

$$y^2 = \tfrac{1}{2}\nu a^2 m^3, \tag{18}$$

which needs modification for m small. Equation (4) gives a vertical boundary with fixed sample size

$$m = 2/(\nu b^2). \tag{19}$$

Finally, let us consider sampling a binomial population in which a proportion θ of individuals possess a certain attribute. Then $v(\theta) = \theta(1-\theta)$. To obtain a specified coefficient of variation b of $\hat{\theta}$, we use the boundary given by equation (4), namely,

$$y = \frac{1}{b^2 + 1/m}, \tag{20}$$

but modified near $m = 0$. For θ small, so that n_0 is large, this is practically the horizontal boundary of 'inverse sampling', which, as already remarked, gives nearly constant coefficient of variation if θ is small. Suppose now that we wish to estimate θ with coefficient of variation b if θ is small, and to estimate $1-\theta$ with the same coefficient of variation if $1-\theta$ is small. We shall meet this requirement if $\ln\{\theta/(1-\theta)\}$ is estimated with constant variance b^2 (assuming, as usual, that the variance is small). Now asymptotically

$$\operatorname{var}[\ln\{\hat{\theta}/(1-\hat{\theta})\}] = [\theta(1-\theta)]^{-2} \operatorname{var}(\hat{\theta}) = [\theta(1-\theta)\, n_0]^{-1}, \tag{21}$$

and this will be equal to b^2 if the boundary satisfies

$$b^2 m \frac{y}{m}\left(1 - \frac{y}{m}\right) = 1,$$

or

$$x + y = b^2 xy, \tag{22}$$

where $x = m - y$. If, therefore, we plot the number of individuals found possessing the attribute as ordinate against the number without as abscissa, sampling proceeds until the rectangular hyperbola (22) is reached. For θ near 0 this is nearly the horizontal straight line $y = b^{-2}$, and for θ near 1 it is nearly the vertical line $x = b^{-2}$.

To sum up, we have considered the estimation of a single unknown parameter. We confine attention to statistics which are the sum Z_n of the observations z_i in some scale of measurement. If we set $\mathrm{E}(z_i) = \theta$, then $\operatorname{var}(z_i)$ can be expressed as a function $v(\theta)$ of θ. We find that the variance of Z_n/n is, for n large, equal to $v(\theta)/n$, whether n is fixed in advance or determined by a single-boundary sequential procedure satisfying conditions (i)–(iii). If we choose a scale of measurement of the observations such that, in samples of fixed size, the unknown parameter is estimated with minimum variance, then the use of the appropriate sequential procedure will secure the desired accuracy of estimation with minimum average sample size.

If there is more than one unknown parameter, and the variance v of the observations z_i is finite but not a known function of the mean θ, v must be estimated as sampling proceeds, and a boundary cannot be specified in advance. Let us suppose that v can be estimated from a sample of m observations by an estimate $\hat{v}$ that is asymptotically normal with variance $O(m^{-1})$ as $m \to \infty$. When, say, half the sampling has been done that will ultimately be seen to be needed, $\hat{v}$ satisfies in probability the relation

$$\hat{v} = v(1 + O(n^{-\frac{1}{2}})), \tag{23}$$

and so the eventual sample size n and estimate $\hat{\theta}$ of θ satisfy

$$\text{var}(\hat{\theta}) \sim \hat{v}/n. \tag{24}$$

Thus the required value of n can be predicted at this half-way stage with asymptotic validity.

I am indebted to Mr D. V. Lindley for helpful criticism. The treatment given above is somewhat heuristic in places, notably in the argument used in arriving at equation (11). It seems to me, however, that the line of approach may prove to be worth pursuing.

REFERENCES

BLACKWELL, D. & GIRSHICK, M. A. (1947). A lower bound for the variance of some unbiased sequential estimates. *Ann. Math. Statist.* **18**, 277.

FINNEY, D. J. (1949). On a method of estimating frequencies. *Biometrika*, **36**, 233.

HALDANE, J. B. S. (1945). On a method of estimating frequencies. *Biometrika*, **33**, 222.

STEIN, C. & WALD, A. (1947). Sequential confidence intervals for the mean of a normal distribution with known variance. *Ann. Math. Statist.* **18**, 427.

TWEEDIE, M. C. K. (1945). Letter in *Nature, Lond.*, **155**, 453.

A historical note on the method of least squares

BY R. L. PLACKETT, *University of Liverpool*

1. The purposes of this note are:

(i) to summarize the justifications by Laplace, Gauss and Markoff of the method of least squares;

(ii) to suggest that Gauss was the first who justified least squares as giving those linear estimates which are unbiased of minimum variance;

(iii) to modernize and extend his proof to cover a general theorem due to Aitken.

It is not my object to provoke controversy, and I have attempted to indicate where a personal opinion is intended.

2. The method of least squares has been in use now for over 150 years. During the nineteenth century the writings of Todhunter (1865), Merriman (1877) and others gave the impression that Laplace (1812–20, collected works 1886) was largely responsible for putting the method on a theoretical basis by means of the calculus of probability, whereas the contribution of Gauss (1821, collected works 1873) was minimized or ignored. Lately, the emphasis has changed, and in recent papers and text-books Markoff (1912) is credited with justifying the method without superfluous assumptions of normality. For these reasons, it seems desirable to disentangle the various justifications proposed and to allot credit in due proportion.

3. In general let $\boldsymbol{\theta}(s \times 1)$ be a vector of unknown parameters, $\mathbf{x}(n \times 1)$ a vector of observations, $\boldsymbol{\epsilon}(n \times 1)$ a vector of errors and $A(n \times s)$ a matrix of known quantities; so that

$$A\boldsymbol{\theta} - \mathbf{x} = \boldsymbol{\epsilon}.$$

Further, suppose that $W(n \times n)$ is a diagonal matrix whose elements are the reciprocals of the error variances. It is required to form an estimate $\boldsymbol{\theta}^*$ of $\boldsymbol{\theta}$. The method of least squares leads to estimates which satisfy

$$A'WA\boldsymbol{\theta}^* = A'W\mathbf{x}.$$

Neither Laplace nor Gauss used matrix notation, but their results can immediately be written in that form.

4. Laplace (1812–20) discusses the method of least squares in Book 2, Chapter 4, and in the first three Supplements. He proves a series of results which are summarized—I hope fairly—in the following:

THEOREM. *Among all $s \times n$ matrices F leading to estimates of the form $FA\theta^* = F\mathbf{x}$, the expected values of the elements of $|\theta^* - \theta|$ are minimized as $n \to \infty$ when $F = \mu A'W$, μ being an arbitrary multiplier.*

The proof is long but runs on these lines: if $\mathbf{u}$ is the error of θ^* then $FA\mathbf{u} = F\boldsymbol{\epsilon}$; Laplace proceeds to determine the joint characteristic function of $F\boldsymbol{\epsilon}$ and deduces that when all errors have the same distribution, symmetrical about zero, $F\boldsymbol{\epsilon}$ has a multivariate normal distribution as $n \to \infty$; whence $\mathbf{u}$ also has a multivariate normal distribution, and the expected values of the elements of $|\mathbf{u}|$ are determined; finally, he shows that $F = \mu A'W$ implies the vanishing of the differential coefficients of these expected values with respect to the elements of F.

In more detail, Laplace first takes $s = 1$ and maximizes the probability that his estimate lies between given limits; he then notes that this is the same as minimizing $\mathscr{E}|\theta^* - \theta|$, and continues to use this criterion when $s = 2$, stating that the result can be extended to greater values of s. In the first Supplement he considers the possibility of a bias in the observations and suggests its removal by introducing an additional parameter whose coefficient is unity in all equations.

5. Gauss presented his justification in 1821. The paper is written in Latin, but a French translation was published by Bertrand in 1855 and the fundamental theorem incorporated in Bertrand's own book of 1888. In the early sections of his paper, Gauss also considers the possibility of bias in his observations and makes it clear that the preferred estimates are those with minimum variance, although of course he does not use this terminology. He begins with errors of differing variance, and by choosing suitable multipliers presents the equations in a form where the errors have the same variance. The proof of the following theorem is in Art. 20; it is implicit that he is seeking unbiased estimates:

THEOREM. *Among all the systems of coefficients $B(s \times n)$ which give $B\boldsymbol{\epsilon} = \boldsymbol{\theta} - \boldsymbol{\theta}\dagger$, the estimate $\boldsymbol{\theta}\dagger$ being independent of $\boldsymbol{\theta}$, those for which the diagonal elements of BB' are minimized are provided by the method of least squares.*

Put $$\boldsymbol{\xi} = A'\boldsymbol{\epsilon} \quad \text{so that} \quad \boldsymbol{\xi} = A'A\boldsymbol{\theta} - A'\mathbf{x}.$$

The solution of these equations is $\boldsymbol{\theta} = \boldsymbol{\theta}^* + D\boldsymbol{\xi}$, and with $DA' = E$ this becomes

$$E\boldsymbol{\epsilon} = \boldsymbol{\theta} - \boldsymbol{\theta}^* \quad \text{so} \quad (\boldsymbol{\theta}^* - \boldsymbol{\theta}\dagger) = (B - E)\boldsymbol{\epsilon}.$$

If this is true for all $\boldsymbol{\theta}$ then $(B - E)A = 0$, i.e. on post-multiplying by D', $(B - E)E' = 0$, which implies $BB' = EE' + (B - E)(B - E)'$. It is now clear that the diagonal elements of BB' are minimized when $B = E$.

6. Matrix notation has been adopted for brevity in the preceding sections, but no matrix theorems have been assumed. Taking for granted the now familiar properties of matrices regarding associative products and inverses, the preceding demonstration can be modernized and shortened.

If $\boldsymbol{\theta}^* = B\mathbf{x}$ is unbiased for all $\boldsymbol{\theta}$, then $BA = I$. With $C = A'A$ it follows that $C^{-1} = BAC^{-1}$, so

$$BB' = (C^{-1}A')(C^{-1}A')' + (B - C^{-1}A')(B - C^{-1}A')',$$

i.e. the diagonal elements of BB' are least when $B = C^{-1}A'$, which is the solution provided by least squares.

7. Markoff (1912) devotes Chapter 7 of his book to the method of least squares. He states that each observation is to be considered as a particular case of many, and as an unbiased estimate of some linear function of the unknown parameters. His determination of unbiased estimates of these parameters having minimum variance is closely followed in the paper by David & Neyman (1938).

8. Aitken (1934) has extended the theorem of Gauss by proving that with a known matrix V of variances and covariances of the observations, the minimum of $(A\boldsymbol{\theta} - \mathbf{x})'V^{-1}(A\boldsymbol{\theta} - \mathbf{x})$ provides estimates $\boldsymbol{\theta}^*$ such that $\boldsymbol{\varphi}^* = P\boldsymbol{\theta}^*$ is an unbiased estimate of $\boldsymbol{\varphi} = P\boldsymbol{\theta}$ with minimum variance. Gauss's method can be used to prove this also.

If $\boldsymbol{\varphi}^* = B\mathbf{x}$ then $BA = P$ and

$$[B]V[P(A'V^{-1}A)^{-1}A'V^{-1}]' = [P(A'V^{-1}A)^{-1}A'V^{-1}]V[P(A'V^{-1}A)^{-1}A'V^{-1}]',$$

consequently

$$BVB' = [P(A'V^{-1}A)^{-1}A'V^{-1}]V[P(A'V^{-1}A)^{-1}A'V^{-1}]' + [B - P(A'V^{-1}A)^{-1}A'V^{-1}]V[B - P(A'V^{-1}A)^{-1}A'V^{-1}]'.$$

If we consider the diagonal elements here, the second term on the right gives a positive definite quadratic form, so minimum variance is attained when

$$B = P(A'V^{-1}A)^{-1}A'V^{-1},$$

the solution given by the method indicated above.

9. It is therefore my opinion that Laplace and Gauss proved theorems which are quite different; that the justification given by Gauss is preferable; and that Markoff, who refers to Gauss's work, may perhaps have clarified assumptions implicit there but proved nothing new. It is evident that Gauss's proof is valid for all values of n, entirely free from any assumption of normality, and capable of immediate development.

REFERENCES

Aitken, A. C. (1934). On least squares and linear combination of observations. *Proc. Roy. Soc. Edinb.* A, **55**, 42–7.

Bertrand, J. (1888). *Calcul des probabilités*. Paris.

David, F. N. & Neyman, J. (1938). Extension of the Markoff theorem on least squares. *Statist. Res. Mem.* **2**, 105–16.

Gauss, C. F. (1855). *Méthode des moindres carrés* (trans. J. Bertrand). Paris.

Gauss, C. F. (1873). Theoria combinationis observationum erroribus minimis obnoxiae. Pars prior. *Werke*, Band **4**. Göttingen.

Laplace, P. S., Marquis de (1886). *Théorie analytique des probabilités*, 3rd edition, Oeuvres, **7**. Paris.

Markoff, A. A. (1912). *Wahrscheinlichkeitsrechnung* (trans. H. Liebmann), 2nd edition. Leipzig and Berlin.

Merriman, M. (1877). A list of writings relating to the method of least squares, with historical and critical notes. *Trans. Conn. Acad. Arts Sci.* **4**, 151–232.

Todhunter, I. (1865). *A History of the Mathematical Theory of Probability*. Macmillan.

The characteristic function of a weighted sum of non-central squares of normal variates subject to s linear restraints

By G. I. Bateman

1. Suppose $x_1, x_2, \ldots, x_n$ are independent normal variables with expectations $a_1, a_2, \ldots, a_n$ respectively and with unit variance, that is to say, we suppose that

$$p(x_j) = \frac{1}{\sqrt{(2\pi)}} \exp[-\tfrac{1}{2}(x_j - a_j)^2], \quad (j = 1, 2, \ldots, n).$$

We consider a weighted non-central sum of squares of the type

$$\psi'^2 = \sum_{j=1}^{n} c_j x_j^2,$$

where the x's are as defined and the c_j $(j = 1, 2, \ldots, n)$ are constants. It is assumed that the x_j are subject to s linear restraints

$$\sum_{j=1}^{n} b_{lj}\, x_j = \rho_l \quad (l = 1, 2, \ldots, s),$$

b_{lj} and ρ_l $(l = 1, \ldots, s;\ j = 1, 2, \ldots, n)$ being given constants. The characteristic function of the joint distribution of $\psi'^2, \rho_1, \rho_2, \ldots, \rho_s$ may be written down immediately. We have

$$\phi(t, t_1, \ldots, t_s) = \prod_{j=1}^{n} \left[\int_{-\infty}^{+\infty} \frac{1}{\sqrt{(2\pi)}} \exp\left\{ -\tfrac{1}{2}(x_j - a_j)^2 + itc_j x_j^2 + \sum_{l=1}^{s} it_l b_{lj} x_j \right\} dx_j \right].$$

The evaluation of the integral is straightforward and we obtain

$$\phi(t, t_1, \ldots, t_s) = \prod_{j=1}^{n} (1 - 2itc_j)^{-\frac{1}{2}} \exp\left(it \sum_{j=1}^{n} \frac{c_j a_j^2}{1 - 2itc_j} - \frac{1}{2} \sum_{l=1}^{s} \sum_{m=1}^{s} A_{lm} t_l t_m + i \sum_{l=1}^{s} B_l t_l \right), \tag{1}$$

where

$$A_{lm} = \sum_{j=1}^{n} \frac{b_{lj} b_{mj}}{1 - 2itc_j} \quad \text{and} \quad B_l = \sum_{j=1}^{n} \frac{a_j b_{lj}}{1 - 2itc_j}. \tag{2}$$

2. Bartlett (1938) considered the case of two variables u and u_1 whose joint characteristic function is $\phi(t, t_1)$, and he showed that the characteristic function of u when u_1 is fixed, which we denote by $\phi(t \mid u_1)$, is given by

$$\phi(t \mid u_1) = \frac{\int_{-\infty}^{+\infty} e^{-itu_1}\phi(t, t_1)\,dt_1}{\int_{-\infty}^{+\infty} e^{-itu_1}\phi(0, t_1)\,dt_1}.$$

This result may be easily, and obviously, extended to the case of $s+1$ variables, $u, u_1, u_2, \ldots, u_s$. Write $p(u, u_1, \ldots, u_s)$ for the joint probability distribution of $u, u_1, \ldots, u_s$ and $\phi(t, t_1, t_2, \ldots, t_s)$ for its characteristic function. It follows that

$$\begin{aligned}\phi(t, t_1, t_2, \ldots, t_s) &= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \exp\left[itu + i\sum_{l=1}^{s} t_l u_l\right] p(u, u_1, \ldots, u_s)\,du\,du_1 \ldots du_s \\ &= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \exp\left[itu + i\sum_{l=1}^{s} t_l u_l\right] p(u \mid u_1, u_2, \ldots, u_s)\,p(u_1, u_2, \ldots, u_s)\,du\,du_1 \ldots du_s.\end{aligned}$$

Writing

$$\phi(t \mid u_1, u_2, \ldots, u_s) = \int_{-\infty}^{+\infty} e^{itu} p(u \mid u_1, u_2, \ldots, u_s)\,du,$$

we have

$$\phi(t, t_1, t_2, \ldots, t_s) = \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \phi(t \mid u_1, u_2, \ldots, u_s)\,p(u_1, u_2, \ldots, u_s) \exp\left[i\sum_{l=1}^{s} t_l u_l\right] du_1 \ldots du_s,$$

whence, using the Fourier transform,

$$\phi(t \mid u_1, u_2, \ldots, u_s)\,p(u_1, u_2, \ldots, u_s) = \frac{1}{(2\pi)^s}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \exp\left[-i\sum_{l=1}^{s} t_l u_l\right] \phi(t, t_1, \ldots, t_s)\,dt_1 \ldots dt_s.$$

It follows that

$$\phi(t \mid u_1 u_2, \ldots, u_s) = \frac{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \exp\left[-i\sum_{l=1}^{s} t_l u_l\right] \phi(t, t_1, t_2, \ldots, t_s)\,dt_1 \ldots dt_s}{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \exp\left[-i\sum_{l=1}^{s} t_l u_l\right] \phi(0, t_1, t_2, \ldots, t_s)\,dt_1 \ldots dt_s}. \tag{3}$$

3. We proceed to apply this result to find the characteristic function of ψ'^2 of § 1. Substituting in the right-hand side of (3) the joint characteristic function of $\psi'^2, \rho_1, \rho_2, \ldots, \rho_s$ as given by (1), we obtain, after some reduction, that the characteristic function of ψ'^2, given $\rho_1, \rho_2, \ldots, \rho_s$, is

$$\begin{aligned}&\phi_{\psi'^2}(t \mid \rho_1, \rho_2, \ldots, \rho_s) \\ &= \prod_{j=1}^{n} (1-2itc_j)^{-\frac{1}{2}} \exp\left[it\sum_{j=1}^{n}\frac{c_j a_j^2}{1-2itc_j}\right] \frac{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \exp\left[i\sum_{l=1}^{s}(B_l-\rho_l)t_l - \frac{1}{2}\sum_{l=1}^{s}\sum_{m=1}^{s} A_{lm}t_l t_m\right] dt_1 \ldots dt_s}{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \exp\left[i\sum_{l=1}^{s}(\beta_l-\rho_l)t_l - \frac{1}{2}\sum_{l=1}^{s}\sum_{m=1}^{s} \alpha_{lm}t_l t_m\right] dt_1 \ldots dt_s},\end{aligned} \tag{4}$$

where

$$\alpha_{lm} = \sum_{j=1}^{n} b_{lj}b_{mj} \quad \text{and} \quad \beta_l = \sum_{j=1}^{n} a_j b_{lj}, \tag{5}$$

and A_{lm} and B_l are as given in (2). The multiple integral in the numerator of (4) can be evaluated and is

$$\frac{(2\pi)^{\frac{1}{2}s}}{\sqrt{|A|}} \exp\left[-\tfrac{1}{2}(B-\rho)'A^{-1}(B-\rho)\right],$$

where A is the matrix $\{A_{lm}\}, l = 1, 2, \ldots, s, m = 1, 2, \ldots, s$, A^{-1} the inverse matrix of A, $|A|$ the determinant associated with A, $(B-\rho)$ the column vector $\{B_l-\rho_l\}, l = 1, 2, \ldots, s$, and $(B-\rho)'$ the corresponding row vector. The denominator of (4) may be evaluated in a similar way and we have finally

$$\begin{aligned}&\phi_{\psi'^2}(t \mid \rho_1, \ldots, \rho_s) \\ &= \prod_{j=1}^{n} (1-2itc_j)^{-\frac{1}{2}} \exp\left[it\sum_{j=1}^{n}\frac{c_j a_j^2}{1-2itc_j}\right] \sqrt{\frac{|\alpha|}{|A|}} \exp\left[-\tfrac{1}{2}(B-\rho)'A^{-1}(B-\rho) + \tfrac{1}{2}(\beta-\rho)'\alpha^{-1}(\beta-\rho)\right],\end{aligned} \tag{6}$$

where α is the matrix $\{\alpha_{lm}\}, l = 1, 2, \ldots, s$, and β the vector $\{\beta_l\}, l = 1, 2, \ldots, s$. The probability law of ψ'^2 for given restraints can be evaluated in certain cases.

4. The distribution of the unweighted sum of non-central squares of normal deviates subject to s orthogonal linear restraints can be obtained as a special case of the foregoing. It has already been derived by Patnaik (1949) using a different method. We put $c_j = 1$ for all j, when

$$\psi'^2 = \sum_{j=1}^{n} x_j^2,$$

where the x's are as previously defined and

$$\mathscr{E}(x_j) = a_j \quad (j = 1, 2, \ldots, n).$$

If the restrictions are orthogonal, i.e. if

$$\sum_{j=1}^{n} b_{lj} b_{mj} = \delta_{lm} = 1 \quad \text{if} \quad l = m,$$
$$= 0 \quad \text{if} \quad l \neq m,$$

then from (2) and (5) we see that

$$A_{lm} = (1-2it)^{-1}\alpha_{lm} = (1-2it)^{-1}\delta_{lm}$$

and

$$B_l = (1-2it)^{-1}\beta_l.$$

Substituting in (4), it follows that

$$\phi_{\sum_{j=1}^{n} x_j^2}(t \mid \rho_1, \ldots, \rho_s) = (1-2it)^{-\frac{1}{2}(n-s)} \exp\left[\frac{it}{1-2it}\sum_{j=1}^{n} a_j^2 - \frac{1}{2(1-2it)}\sum_{l=1}^{s}(\beta_l - \rho_l(1-2it))^2 + \frac{1}{2}\sum_{l=1}^{s}(\beta_l - \rho_l)^2\right]$$

$$= (1-2it)^{-\frac{1}{2}(n-s)} \exp\left[\frac{it\lambda}{1-2it} + it\sum_{l=1}^{s}\rho_l^2\right], \tag{7}$$

where

$$\lambda = \sum_{j=1}^{n} a_j^2 - \sum_{l=1}^{s}\left(\sum_{j=1}^{n} a_j b_{lj}\right)^2. \tag{8}$$

We recall that the characteristic function of the non-central χ^2, referred to as χ'^2, is

$$\phi_{\chi'^2}(t) = (1-2it)^{-\frac{1}{2}n} \exp\left[\frac{\lambda it}{1-2it}\right].$$

Hence, using the uniqueness property of the characteristic function and the fact that there is a (1, 1) correspondence between the characteristic function of a variable and its probability law, it follows that when the x_j's are subject to s orthogonal linear restraints,

$$\sum_{j=1}^{n} x_j^2 - \sum_{l=1}^{s}\rho_l^2$$

is distributed as χ'^2 with $(n-s)$ degrees of freedom and with parameter λ as given by (8).

REFERENCES

BARTLETT, M. S. (1938). *J. Lond. Math. Soc.* **13**, 62.
PATNAIK, P. B. (1949). *Biometrika*, **36**, 202.

Intra-class rank correlation

By J. W. WHITFIELD

Psychological Laboratory, University of Cambridge

Given a number of pairs of items each of which is measured according to some quality, the normal procedure to discover whether the arrangement in pairs bears any relation to the quality measured is to calculate the intra-class correlation coefficient. As an example we may consider pairs of brothers and a measure of their stature. If we wish to know whether there is any correlation between older brothers and younger brothers on stature, we calculate the product-moment correlation in the ordinary way. But if the more general question whether brothers tend to have similar stature is asked, there are no grounds upon which we can differentiate between each member of a pair so as to separate them into the two arrays for correlation. The mean correlation for all possible arrangements is required. This is the intra-class correlation which can be calculated without performing the numerous individual product-moment correlations. With quantitative data the procedure is not limited to paired data, but may be performed with arrangements of varying sized groups.

The same type of problem can arise with ranked data. Consider eight pairs of brothers, ranked upon some quality which is not amenable to quantitative measure:

	Rank values
First pair:	1 and 6
Second pair:	5 and 9
Third pair:	7 and 8
Fourth pair:	11 and 13
Fifth pair:	2 and 4
Sixth pair:	15 and 16
Seventh pair:	10 and 14
Eighth pair:	2 and 12

There are 256 possible arrangements for correlation, although only 128 need actually be calculated. These are distributed (using Kendall's τ coefficient of rank correlation—Kendall, 1948) as follows:

Frequency	S value
8	+18
32	+16
40	+14
16	+12
8	+10
16	+ 8
8	+ 6
128	+1664

Hence the mean S is +13, and the mean τ is +0·4643.

The computation of S for all possible arrangements would be excessively tedious for all data except those involving very few observations. The following alternative computation, though restricted to the case of paired data only, is much simpler. It can probably be extended to include family groupings other than two.

Rearrange the data in the example in order of the higher rank in each pair—i.e. first the pair with the highest ranked individual in it, next the pair with the highest ranked individual from the remainder, etc., giving

(1, 6) (2, 4) (3, 12) (5, 9) (7, 8) (10, 14) (11, 13) (15, 16).

Calculate S for this order (making no comparisons within pairs). For example, the first number, 1, has 14 numbers on its right greater than itself, ignoring the number 6 which is bracketed with it, and the contribution to the score is 14. The second number, 6, has 10 greater numbers on its right and the contribution is 10. The third number, 2, has 13 greater numbers, ignoring the 4 which is bracketed with it; and so on. The positive score obtained in this way is

$$14+10+13+11+10+3+8+6+6+6+4+2+2+2 = 97.$$

The maximum score is 112 and hence the score of the negative contributions is $112-97 = 15$. The total score is thus $97-15 = 82$. The maximum value S can attain is $\frac{1}{2}n(n-2)$, where n is the number of individuals. The minimum value of S is zero. This follows from a consideration of the extreme case of first paired with nth, second with $(n-1)$th, and so on, where the S contribution from the higher ranked individuals in the pairs is $\frac{1}{2}n(\frac{1}{2}n-1)$ and the S contribution from the lower ranked members is $-\frac{1}{2}n(\frac{1}{2}n-1)$. Furthermore, as will be seen from the sampling distribution, the mean value lies midway between zero and $\frac{1}{2}n(n-2)$, and the sampling distribution is symmetrical about the mean. These qualities of mean and range are true only for families of two members each. Taking the mean as origin, we can define

$$S_p = S - \frac{n(n-2)}{4}$$

(using the suffix p to indicate paired data) and

$$r_p = \frac{4S_p}{n(n-2)}.$$

In the example $$S_p = 82-\frac{16.14}{4} = +26, \quad r_p = +\frac{4.26}{16.14} = +0{\cdot}4643,$$

which agrees with the value of mean r obtained from considering all possible arrangements for correlation.

Sampling Distribution of S_p

The sampling distribution of S_p has been calculated for $n = 6, 8, 10, \ldots, 20$, and the tables are given at the end of this paper. A summation feature makes the calculation easy. The simplest case is that of $n = 4$, for which we have:

Possible orders	S	S_p
(1, 2) (3, 4)	4	2
(1, 3) (2, 4)	2	0
(1, 4) (2, 3)	0	−2

For $n = 6$, we have S (or S_p) composed of the internal contribution of the second and third pairs, distributed as for $n = 4$, together with the contribution of the first pair with respect to the second and third pairs.

The pair (1, 2) will occur with S values $8+4$, $8+2$, $8+0$;
The pair (1, 3) will occur with S values $6+4$, $6+2$, $6+0$;
The pair (1, 4) will occur with S values $4+4$, $4+2$, $4+0$;
The pair (1, 5) will occur with S values $2+4$, $2+2$, $2+0$;
The pair (1, 6) will occur with S values $0+4$, $0+2$, $0+0$;

or, expressed as a frequency table

S	S_p	First pair					Frequency
		(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)	
12	6	1	.	.	.	.	1
10	4	1	1	.	.	.	2
8	2	1	1	1	.	.	3
6	0	.	1	1	1	.	3
4	−2	.	.	1	1	1	3
2	−4	.	.	.	1	1	2
0	−6	.	.	.	.	1	1

The sampling distribution for n is, therefore, the aggregate of $(n-1)$ sampling distributions for $(n-2)$, distributed with means at $+(n-2)$, $(n-4)$, ..., $-(n-4)$, $-(n-2)S_p$. This feature demonstrates the symmetry of the distribution, and also makes for a simple calculation of the variance.

Using $c_p = \frac{1}{2}S_p$, we have the sum of squares of c_{p_n} about zero in two parts (a) $(n-1)$ times the sum of squares for the distribution of $c_{p_{n-1}}$, together with (b) the contribution due to the displacement of these subgroups from zero, i.e.

$$\text{Sum of squares } c_{p_n} = (n-1)\,(\text{sum of squares } c_{p_{n-1}}) + 2\{(n-3)(n-5)\ldots 1\}\left\{\left(\frac{n-2}{2}\right)^2+\left(\frac{n-4}{2}\right)^2+\ldots(1)^2\right\}. \qquad (1)$$

Expanding (1) we get

$$2\Bigg[\{(n-3)(n-5)\ldots 1\}\left\{\left(\frac{n-2}{2}\right)^2+\left(\frac{n-4}{2}\right)^2+\ldots(1)^2\right\} + \{(n-1)(n-5)(n-7)\ldots 1\}\left\{\left(\frac{n-4}{2}\right)^2+\left(\frac{n-6}{2}\right)^2+\ldots(1)^2\right\} + \{(n-1)(n-3)(n-7)(n-9)\ldots(1)\} \times\left\{\left(\frac{n-6}{2}\right)^2+\left(\frac{n-8}{2}\right)^2+\ldots(1)^2\right\}+\ldots\Bigg].$$

Dividing by $\{(n-1)(n-3)(n-5)\ldots 1\}$ to obtain the variance

$$\operatorname{var} c_{p_n} = 2\left[\frac{\left\{\left(\frac{n-2}{2}\right)^2+\left(\frac{n-4}{2}\right)^2+\ldots(1)^2\right\}}{n-1}+\frac{\left\{\left(\frac{n-4}{2}\right)^2+\left(\frac{n-6}{2}\right)^2+\ldots(1)^2\right\}}{n-3} + \frac{\left\{\left(\frac{n-6}{2}\right)^2+\left(\frac{n-8}{2}\right)^2+\ldots(1)^2\right\}}{n-5}+\ldots\right]. \qquad (2)$$

Substituting $l = \frac{1}{2}n$ in (2) we find

$$\begin{aligned}\operatorname{var} c_{p_n} &= 2\left[\frac{(l-1)^2+(l-2)^2+\ldots(1)^2}{2l-1}+\frac{(l-2)^2+(l-3)^2+\ldots(1)^2}{2l-3}+\frac{(l-3)^2+(l-4)^2+\ldots(1)^2}{2l-5}+\ldots\right]\\ &= \frac{2}{6}\left[\frac{l(l-1)(2l-1)}{2l-1}+\frac{(l-1)(l-2)(2l-3)}{2l-3}+\frac{(l-2)(l-3)(2l-5)}{2l-5}+\ldots\right]\\ &= \tfrac{1}{3}\{l(l-1)+(l-1)(l-2)+(l-2)(l-3)+\ldots\}\\ &= \frac{1}{3}\left\{\frac{l(l+1)(l-1)}{3}\right\} = \frac{l^3-l}{9}\end{aligned}$$

Hence

$$\operatorname{var} c_{p_n} = \frac{n^3-4n}{72} \quad \text{or} \quad \frac{n(n-2)(n+2)}{72}$$

$$\operatorname{var} S_{p_n} = \frac{n^3-4n}{18} \quad \text{or} \quad \frac{n(n-2)(n+2)}{18}.$$

It is necessary to consider whether the distribution approaches normality so that a normal deviate test can be applied to an obtained value of S_p. As the distribution is symmetrical, the third moment and β_1 are zero. Fourth moments have been computed for the distributions $n = 4$ to $n = 20$, and β_2 coefficients calculated. At low values of n the distribution is distinctly platykurtic, but the β_2 value approaches 3 as n increases. The values of β_2 are given at the foot of the tables at the end of this note.

The fourth moment of S_p can be calculated in a manner similar to that used above to derive the variance. The algebra can, however, be avoided if we note that μ_4 must be a sextic in n and fit such a sextic to the first seven values of the fourth moment, which are as follows:

n	μ_4
4	10·6667
6	247·4667
8	1719·4667
10	7244·8000
12	22892·8000
14	59852·8000
16	136729·6000

Probability that a given value of S_p will be attained or exceeded by chance (single tail only, distribution symmetrical about zero)

S_p	$n=6$	$n=8$	$n=10$	$n=12$	$n=14$	$n=16$	$n=18$	$n=20$	S_p
0	0·50000	0·50000	0·50000	0·50000	0·50000	0·50000	0·50000	0·50000	0
2	0·40000	0·42857	0·44868	0·46080	0·46875	0·47432	0·47842	0·48153	2
4	0·20000	0·29524	0·34921	0·38374	0·40693	0·42336	0·43549	0·44473	4
6	0·06667	0·18095	0·25820	0·31063	0·34717	0·37356	0·39326	0·40838	6
8	—	0·09524	0·17989	0·24367	0·29069	0·32564	0·35217	0·37276	8
10	—	0·03810	0·11640	0·18461	0·23855	0·28025	0·31264	0·33813	10
12	—	0·00952	0·06878	0·13499	0·19156	0·23794	0·27502	0·30475	12
14	—	—	0·03598	0·09370	0·15023	0·19913	0·23964	0·27283	14
16	—	—	0·01587	0·06195	0·11483	0·16412	0·20673	0·24257	16
18	—	—	0·00529	0·03848	0·08532	0·13309	0·17649	0·21412	18
20	—	—	0·00106	0·02213	0·06143	0·10606	0·14903	0·18760	20
22	—	—	—	0·01154	0·04268	0·08296	0·12440	0·16309	22
24	—	—	—	0·00529	0·02843	0·06359	0·10258	0·14065	24
26	—	—	—	0·00202	0·01814	0·04769	0·08352	0·12028	26
28	—	—	—	0·00058	0·01093	0·03492	0·06708	0·10196	28
30	—	—	—	0·00010	0·00616	0·02490	0·05310	0·08565	30
32	—	—	—	—	0·00320	0·01725	0·04140	0·07127	32
34	—	—	—	—	0·00150	0·01156	0·03175	0·05871	34
36	—	—	—	—	0·00061	0·00747	0·02392	0·04786	36
38	—	—	—	—	0·00021	0·00462	0·01768	0·03859	38
40	—	—	—	—	0·00005	0·00272	0·01280	0·03076	40
42	—	—	—	—	0·00001	0·00151	0·00906	0·02421	42
44	—	—	—	—	—	0·00078	0·00626	0·01882	44
46	—	—	—	—	—	0·00037	0·00420	0·01442	46
48	—	—	—	—	—	0·00016	0·00274	0·01089	48
50	—	—	—	—	—	0·00006	0·00172	0·00810	50
52	—	—	—	—	—	0·00002	0·00104	0·00592	52
54	—	—	—	—	—	$0·0^{5}4$	0·00060	0·00425	54
56	—	—	—	—	—	$0·0^{6}5$	0·00033	0·00299	56
58	—	—	—	—	—	—	0·00017	0·00206	58
60	—	—	—	—	—	—	0·00008	0·00138	60
62	—	—	—	—	—	—	0·00004	0·00091	62
64	—	—	—	—	—	—	0·00001	0·00058	64
66	—	—	—	—	—	—	$0·0^{5}5$	0·00035	66
68	—	—	—	—	—	—	$0·0^{5}1$	0·00021	68
70	—	—	—	—	—	—	$0·0^{6}3$	0·00012	70
72	—	—	—	—	—	—	$0·0^{7}3$	0·00007	72
74	—	—	—	—	—	—	—	0·00003	74
76	—	—	—	—	—	—	—	0·00002	76
78	—	—	—	—	—	—	—	0·00001	78
80	—	—	—	—	—	—	—	$0·0^{5}3$	80
82	—	—	—	—	—	—	—	$0·0^{5}1$	82
84	—	—	—	—	—	—	—	$0·0^{6}3$	84
86	—	—	—	—	—	—	—	$0·0^{7}8$	86
88	—	—	—	—	—	—	—	$0·0^{7}2$	88
90	—	—	—	—	—	—	—	$0·0^{8}2$	90
β_2	2·17	2·42	2·55	2·63	2·68	2·72	2·76	2·80	β_2

From this we get
$$\mu_4 = \frac{n^6}{108} - \frac{n^5}{75} - \frac{2n^4}{27} + \frac{2n^3}{45} + \frac{4n^2}{27} + \frac{8n}{225},$$

and hence
$$\beta_2 \sim 3 - \frac{4{\cdot}32}{n}.$$

Thus β_2 will approach the value 3 as n increases, and hence for reasonable values of n the normal deviate test can be employed.

Uses of method

It seems to the writer that there are several situations in experimental psychology where this method will be of use, in addition to the type of situation used in the example. One such possibility concerns the problem of relating preference judgements to particular qualities of the objects judged. As an example, in preference judgements about different types of chair, height of seat from the ground may be an important factor, but the optimum may vary from individual to individual. If various types of chairs were chosen so that they could be considered as a number of pairs with regard to this quality (a kind of ranking replication), the preference judgements could be treated in this fashion independently for each judge, to discover whether or not seat height is an important criterion in the preferences.

REFERENCE

Kendall, M. G. (1948). *Rank Correlation Methods.* London: Griffin and Co.

A note on non-normal correlation

By J. B. S. HALDANE

The product-moment correlation ρ is frequently estimated for two variates which are not normally distributed. There are, however, no general expressions for the effect of this non-normal distribution on the precision of the estimate of ρ. They may be obtained in one special case which is of biological importance. Suppose X and Y are two correlated variates. Then if

$$X = a + \frac{\sigma}{\sqrt{2}}[(1+\rho)^{\frac{1}{2}}x + (1-\rho)^{\frac{1}{2}}y], \quad Y = b + \frac{\tau}{\sqrt{2}}[(1+\rho)^{\frac{1}{2}}x - (1-\rho)^{\frac{1}{2}}y],$$

and further if $\bar{x} = \bar{y} = 0$, $\bar{x}^2 = \bar{y}^2 = 1$, and x and y are independent, the variance of X is σ^2, that of Y is τ^2, and their covariance is $\rho\sigma\tau$, regardless of the distributions of x and y. Hence the correlation of X and Y is ρ. If x and y are normally distributed the correlation is of course normal. Now in biological statistics X and Y may be measurements of two organs in the same individual, or of their logarithms. x depends on the sum of causes which affect X and Y alike, y on the sum of causes which affect them oppositely. For example, in any series of specimens, not all of which are fully grown, x will increase with age up to a certain point; and in a population containing a minority of juvenile members the distribution of x will probably be negatively skew. But y may be quite independent of age if the variability of the organs measured is uncorrelated with age, and may well be normally distributed when x is not.

Let κ_{rs} be the cumulants of the joint distribution of x and y. Then since they are independent, $\kappa_{rs} = 0$ unless r or $s = 0$, $\kappa_{10} = \kappa_{01} = 0$, $\kappa_{20} = \kappa_{02} = 1$; and let $\kappa_{30} = \gamma_1$, $\kappa_{03} = \gamma_1'$, $\kappa_{40} = \gamma_2$, $\kappa_{04} = \gamma_2'$, etc., these being measures of the deviations from normality of the distributions of x and y.

Our estimate of ρ on a sample of n members is thus

$$r = \frac{n\Sigma X_r Y_r - \Sigma X_r \Sigma Y_r}{[n\Sigma X_r^2 - (\Sigma X_r)^2]^{\frac{1}{2}}[n\Sigma Y_r^2 - (\Sigma Y_r)^2]^{\frac{1}{2}}}$$

$$= [n(1+\rho)\Sigma x_r^2 - n(1-\rho)\Sigma y_r^2 - (1+\rho)(\Sigma x_r)^2 + (1-\rho)(\Sigma y_r)^2]$$
$$\times [n(1+\rho)\Sigma x_r^2 + n(1-\rho)\Sigma y_r^2 + 2n(1-\rho^2)^{\frac{1}{2}}\Sigma x_r y_r - (1+\rho)(\Sigma x_r)^2 - (1-\rho)(\Sigma y_r)^2 - 2(1-\rho^2)^{\frac{1}{2}}\Sigma x_r \Sigma y_r]^{-\frac{1}{2}}$$
$$\times [n(1+\rho)\Sigma x_r^2 + n(1-\rho)\Sigma y_r^2 - 2n(1-\rho^2)^{\frac{1}{2}}\Sigma x_r y_r - (1+\rho)(\Sigma x_r)^2 - (1+\rho)(\Sigma y_r)^2 + 2(1-\rho^2)^{\frac{1}{2}}\Sigma x_r \Sigma y_r]^{-\frac{1}{2}}.$$

So
$$r^2 = \frac{[(1+\rho)\{n\Sigma x_r^2 - (\Sigma x_r)^2\} - (1-\rho)\{n\Sigma y_r^2 - (\Sigma y_r)^2\}]^2}{[(1+\rho)\{n\Sigma x_r^2 - (\Sigma x_r)^2\} + (1-\rho)\{n\Sigma y_r^2 - (\Sigma y_r)^2\}]^2 - 4(1-\rho^2)(n\Sigma x_r y_r - \Sigma x_r \Sigma y_r)^2}$$
$$= \frac{[(1+\rho)k_{20} - (1-\rho)k_{02}]^2}{[(1+\rho)k_{20} + (1-\rho)k_{02}]^2 - 4(1-\rho^2)k_{11}^2}, \qquad (1)$$

where k_{rs} is the unbiased estimate of κ_{rs} from the moments of the variates in the sample. For example, $k_{20} = \dfrac{n\Sigma x_r^2 - (\Sigma x_r)^2}{n(n-1)}$. We can now ask how the mean value of r^2 will be affected by deviations of the distributions of x and y from normality. $\overline{k_{20}^2}$ exceeds $(\overline{k_{20}})^2$, or unity, by the sampling variance of k_{20} which is $\dfrac{2\kappa_{20}^2}{n-1} + \dfrac{\kappa_{40}}{n}$, or $\dfrac{2}{n-1} + \dfrac{\gamma_2}{n}$. The effect of non-normality in the distribution of x is therefore to increase the mean value of k_{20}^2 by γ_2/n. Similarly, $\overline{k_{02}^2}$ is increased by γ_2'/n. $\overline{k_{20}k_{02}}$ is not increased, since x and y are independent. k_{11}^2 does not include terms with zero suffixes, so it is also unaltered. In fact, both numerator and denominator of (1) are increased by $n^{-1}[(1+\rho)^2\gamma_2 + (1-\rho)^2\gamma_2']$.

We cannot calculate the variance of r directly from (1) since $\bar{r}$ differs from ρ by a quantity of order n^{-1}. But since both the numerator and denominator are increased by

$$(1+\rho)^2\frac{\kappa_{40}}{n} + (1-\rho)^2\frac{\kappa_{04}}{n} \quad \text{or} \quad n^{-1}[(1+\rho)^2\gamma_2 + (1-\rho)^2\gamma_2']$$

above the values found when the distributions of x and y are normal, we have in the normal case

$$r_0^2 = \frac{4\rho^2 + 4n^{-1}P}{4 + 4n^{-1}Q},$$

where P and Q are independent of n to order n^{-2}, and in general

$$\overline{r^2} = \frac{4\rho^2 + n^{-1}[4P + (1+\rho)^2\gamma_2 + (1-\rho)^2\gamma_2']}{4 + n^{-1}[4Q + (1+\rho)^2\gamma_2 + (1-\rho)^2\gamma_2']}.$$

So
$$\overline{r^2} - \overline{r_0^2} = \frac{1-\rho^2}{4n}[(1+\rho)^2\gamma_2 + (1-\rho)^2\gamma_2'] + O(n^{-2}).$$

The variance of r is therefore increased by this quantity. The precision of the estimate of ρ does not therefore depend on the skewness of the distributions of x and y, provided they are mesokurtic. And since

$$\gamma_1 = \sqrt{2}(1+\rho)^{-\frac{1}{2}}[\gamma_1(X) + \gamma_1(Y)], \quad \gamma_1' = \sqrt{2}(1-\rho)^{-\frac{1}{2}}[\gamma_1(X) - \gamma_1(Y)],$$

it follows that skew distribution of X and Y will not affect the precision of r. On the other hand, the distributions of X and Y have the same value of γ_2 or $\beta_2 - 3$, namely,

$$\Gamma_2 = \tfrac{1}{4}[(1+\rho)^2\gamma_1 + (1-\rho)^2\gamma_2].$$

Hence
$$\operatorname{var}(r) = \frac{1-\rho^2}{n}(1-\rho^2+\Gamma_2) + O(n^{-2}). \tag{2}$$

If we employ Fisher's transformation $z = \frac{1}{2}\log\left(\dfrac{1+r}{1-r}\right)$, we find

$$\operatorname{var}(z) = n^{-1}\left(1 + \frac{\Gamma_2}{1-\rho^2}\right) + O(n^{-2}). \tag{3}$$

The variance of z is thus no longer almost independent of ρ. But the precision of r is increased if the distributions of X and Y are platykurtic, and decreased if they are leptokurtic. Clearly, however, (2) and (3) are inapplicable when $|\rho|$ is near unity, terms of at least order n^{-2} being required.

On empirical grounds, E. S. Pearson (1931, 1932) stated that 'the normal bivariate surface may be distorted and mutilated to a remarkable degree without affecting the frequency distribution of r'. This would seem to be true when $|\rho|$ is not near unity. It is also true for 'mutilations' which affect skewness without doing a great deal to kurtosis. However, when correlation is high it would seem that a relatively slight change in kurtosis may have a large effect on the variance of r.

It is possible that the formula (1) might serve as a basis for a new development of the theory of the distribution of r in the normal case, and further information could certainly be obtained from it concerning the more general case here considered. In the most general case x and y, though they have a zero coefficient of correlation, are not independent, so such cumulants as κ_{22} would not in general be zero, and it is doubtful whether the method would be of value. On the other hand, if the distributions of X and Y, though having different values of β_1, have insignificantly different values of β_2, equations (2) or (3) may be used with some confidence.

I have to thank Mr K. A. Kermack for useful criticism.

REFERENCE

PEARSON, E. S. (1931, 1932). The test of significance for the correlation coefficient. *J. Amer. Statist. Soc.* **26**, 128; **28**, 424.

REVIEWS

Probability Theory for Statistical Methods. By F. N. DAVID. ix+230 pp. Cambridge University Press. 1949. Price 15*s*.

This book deals with the simpler parts of probability theory, in so far as they are relevant to common statistical techniques. It is intended primarily for students who have reached intermediate standard in mathematics, but it will be found useful by all who are interested in statistical methods, as it gathers together a number of useful results which otherwise are very much scattered throughout the literature.

The book begins with a discussion of the definition of probability; after considering various viewpoints, the author decides on a definition which is adequate for the purposes of the book (except that strictly speaking in the form given it applies only to *rational* probabilities, whereas when the normal distribution is being considered we shall need irrational ones too). After this the author goes on to consider the binomial distribution. Three ways of summing the tails of a binomial are discussed: the exact method, using the incomplete beta function; a not very well known but very convenient continued fraction due to Markoff and Müller; and also the usual normal approximation.

Various generalizations of the binomial distribution are then discussed: Poisson's limit, the negative binomial, Neyman's contagious distribution, and, in later chapters, the multinomial distribution, and Poisson's and Lexis's forms of the binomial for heterogeneous data. Methods are given for fitting most of these distributions to observed data.

A chapter is devoted to Bayes's Theorem, pointing out the conditions under which it may be validly applied, and also to confidence intervals, and their uses. There follows a chapter on simple genetical applications of the theory, largely replacing the usual sequence of problems on drawing black and white balls from urns, throwing dice, etc. This seems a very welcome development, provided that it is accompanied by a sympathetic understanding of at least the elements of genetics, and genes are not going to be regarded in future simply as objects conveniently provided by nature for the construction of new types of examination questions. Incidentally, Dr David is over-cautious in her discussion of random mating, saying that it is 'rarely met with in practice', whereas in fact with genes such as the blood-group genes it can be shown to hold with considerable accuracy.

Three chapters deal with random variables, their distributions, 'expected' values, and moments, and with the law of large numbers. The simpler moments of the sample mean and sample variance are worked out in some detail, both for infinite and (in the case of the mean) for finite populations.

Two chapters on estimation follow. Here the estimates are practically restricted to linear unbiased estimates; and the Generalized Markoff Theorem provides a method of obtaining these estimates—a proof of the theorem is given for the case of two parameters. One important application of this theorem is to linear regression, and another to the problem of efficient sampling of human and other populations. Here the methods of stratified and restricted stratified sampling are explained and discussed in some detail.

The final three chapters are concerned with characteristic functions and their elementary properties. Here we are concerned with finding the characteristic function of a known distribution, and with the converse, with the connexion between characteristic functions, moments, and cumulants, and with the central limit theorem.

The style of the book is simple, straightforward, and lucid—refreshingly so; and there are few errors, mostly trivial ones. The only two places where the exposition is not quite up to the standard of the rest of the book seem to be the 'problem of points', p. 26, where the condition $r<t$ might be brought in earlier, and the calculation on pp. 96–7, where the sentence 'we may assume therefore that $A_1 = \mathbf{Rr}$' gives rise to what seem at first sight rather doubtful calculations of the probabilities of various genotypes for A_2; if, however, we say instead 'let A_1 be the parent who passed on the **r** gene to B_2' then these calculations are completely justified. Incidentally the phrase 'for example, colour-blindness' which occurs in the examination question quoted here seems to be singularly unfortunate, as colour-blindness is not inherited in this manner, nor has it a gene frequency as low as 0·001. Other omissions which may cause some difficulty occur on pp. 15–16, where the product law is used before it has been proved; in chapter XV, where the fact that two different distributions cannot have the same characteristic function is frequently used but nowhere stated; and on p. 214, where the independence of the grouping error is tacitly assumed. It is a pity that means and moments are not defined on p. 31, and their simpler relations proved, since apart from this omission the book is self-contained. Small errors occur on p. 55, line 6, where the numerator should be $|k-np|-\frac{1}{2}$; p. 59, line 8, where the inequality should read $\ldots \leqslant \left(1-\frac{\beta}{2n}\right)^2$; p. 63, second formula,

read $(2x+t)$ for $(2n+t)$ in square brackets; p. 96, line 9, for 'their offspring' read 'a single one of their offspring' (this was expressed badly in the original question); p. 107, line 9, for 105 read 10^5; p. 126, line 10, for $(n/n-1)^{\frac{1}{2}}$ read $[n/(n-1)]^{\frac{1}{2}}$; p. 128, line 4 from foot, for 'central difference' read 'linear difference'; p. 147, Markoff's Lemma should read '$\ldots \leqslant 1/t^2$', and similarly on p. 148, the Bienaimé-Tchebycheff inequality '$\ldots \geqslant 1-1/t^2$'; p. 165, section (iii), 'the n equations for $\mathscr{E}(x_t)$'. Finally, on p. 205, the definition of boundedness does not seem to be completely happily phrased.

There are also a few points on which the reviewer would like to express a personal opinion, although perhaps not everyone would agree. The use of $(q+p)^n$ instead of (say) $(q+tp)^n$ as a generating function (as on p. 30) does not seem very happy, since obviously $(q+p)^n = 1$. The consistent use of 'standard error' for 'standard deviation' seems also a little odd. The usual practice for geneticists now is to concentrate attention on gene frequencies, rather than the genotype frequencies considered in chapter VIII; although this would probably simplify the calculations, the point is not very important in the context. The French spelling 'Tchebycheff', though in common use, is irritating, but can perhaps be blamed on the man himself, although he was really 'Chebyshôf'. And finally, the notation $\Gamma(x+1)$ for $x!$, often used when x is not integral, seems merely designed to make the formulae unnecessarily complicated.

Summing up, we may say that the book would be a most useful addition to the library of any statistician or student of statistics.

CEDRIC A. B. SMITH

The Fundamentals of Statistics. By T. L. Kelley. 755 pp. Harvard University Press, and Oxford University Press. 1947. Price 55*s*.

This book is an introductory text to statistical theory. The author aims at setting out the logical processes which are involved in the application of any piece of statistical technique and generally succeeds with admirable clarity. Many who are interested in what statistical methods are about but do not care for mathematics may read this text with profit, for the mathematical theory is kept to a minimum in the first 200 pages. It is perhaps open to question whether Prof. Kelley achieves the clarity with his mathematical analysis that he does with his exposition of general principles. The reviewer found his notation cumbersome in places, and it is likely that a student coming fresh to the subject will very often find himself in difficulties. For example, the section on sequential analysis gives the outline only of the mathematical analysis. Reference is given to other publications whereby the student may supplement this analysis, but for the elementary student these other publications will prove too difficult.

On the whole it would be fair to say that this book is a useful addition to statistical text-books, although it will never take the place of such old and tried classics as Yule and Kendall's *Elementary Theory of Statistics*, for the student beginning the study of statistical methods. Biologists, medicals, etc., will probably find it useful to read before attempting R. A. Fisher's *Statistical Methods for Research Workers*.

F. N. DAVID

(*All Rights reserved*)

BIOMETRIKA. Vol. XXXVI, Parts III and IV

CONTENTS

	PAGE
The estimation of the parameters of tolerance distributions. By D. J. Finney	239–256
An overlap problem arising in particle counting. By P. Armitage	257–266
Tables of autoregressive series. By M. G. Kendall	267–289
Tables for use in comparisons whose accuracy involves two variances, separately estimated. By Alice A. Aspin. With an Appendix by B. L. Welch	290–296
Bivariate distributions based on simple translation systems. By N. L. Johnson	297–304
A test for randomness in a sequence of two alternatives involving a 2×2 table. By P. G. Moore	305–316
A general distribution theory for a class of likelihood criteria. By G. E. P. Box	317–346
Note on approximations to the power function of the '2×2 comparative trial'. By G. P. Sillitto	347–352
The distribution of 'Student's' t in random samples of any size drawn from non-normal universes. By A. K. Gayen	353–369
The combination of probabilities arising from data in discrete distributions. By H. O. Lancaster	370–382
Note on the application of Fisher's k-statistics. By F. N. David	383–393
The moments of the z and F distributions. By F. N. David	394–403
The method of frequency-moments and its application to type VII populations. By Herbert S. Sichel	404–425
On the use of Student's t-test in an asymmetrical population. By S. G. Ghurye	426–430
Tables of symmetric functions—Part I. By F. N. David and M. G. Kendall	431–449
MISCELLANEA	
On the efficiency of the method of moments and Neyman's type A distribution. By L. R. Shenton	450–454
Large-sample theory of sequential estimation. By F. J. Anscombe	455–458
A historical note on the method of least squares. By R. L. Plackett	458–460
The characteristic function of a weighted sum of non-central squares of normal variates subject to s linear restraints. By G. I. Bateman	460–462
Intra-class rank correlation. By J. W. Whitfield	463–467
A note on non-normal correlation. By J. B. S. Haldane	467–468
REVIEWS	
F. N. David's 'Probability Theory for Statistical Methods'	469–470
T. L. Kelley's 'The Fundamentals of Statistics'	470

A volume of *Biometrika* containing about 400 pages, with plates and tables, will be published annually in two half-yearly issues.

Papers for publication should either be sent to

PROFESSOR E. S. PEARSON, Department of Statistics, University College, London, W.C. 1,

or if more convenient to

Dr John Wishart, Statistical Laboratory, St Andrew's Hill, Cambridge.

Professor M. G. Kendall, Cudham Court, Cudham, near Knockholt, Kent.

It is a condition of publication in *Biometrika* that the paper shall not already have been issued elsewhere, and will not be reprinted without leave of the Editors.

Contributors receive 25 copies of their papers free. Joint authors 15 copies each.

The subscription price, *payable in advance*, is Inland 45*s. net* per volume and Abroad 54*s. net* (including packing and postage). Recent volumes may still be obtained at the wrapper price; this is 64*s.* inland, including postage. Cheques must be made payable to *Biometrika*, crossed "*a/c Biometrika Trust*" and sent to The Secretary, *Biometrika Office*, Department of Statistics, University College, London, W.C. 1, to whom all orders for series, single copies and offprints should be addressed. All foreign cheques must be drawn in sterling and on a Bank having a London Agency.

First Printed in Great Britain at the University Press, Cambridge
Reprinted by offset-litho by Percy Lund Humphries & Co., Ltd., Bradford